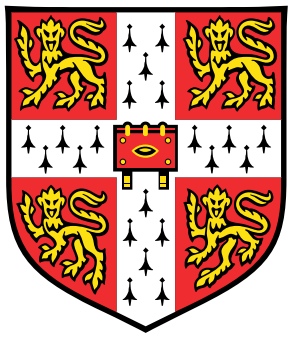
Exploring the diversity of epidermal pavement cell shapes



Róza Villő Vőfély

Department of Plant Sciences
University of Cambridge

—

Trinity College

This dissertation is submitted for the degree of
Doctor of Philosophy

September 2017
Corrections submitted August 2018

Abstract

Pavement cells in the leaf epidermis can exhibit a wide variety of shapes. Undulations in the anticlinal cell wall are a particularly interesting feature present in many model species: for example, *Arabidopsis thaliana*, tobacco and maize; but absent in others, like *Brachypodium distachyon*. Epidermal cell patterning is qualitatively similar between individuals of the same species but different species can have wildly different shapes and there is no comprehensive data set available for epidermal cell shapes. Moreover, the cell shape changes during development this has only been quantitatively studied in *Arabidopsis*.

This work presents and analyses three sets of cell shape data. The first part examines the morphogenesis of the curiously shaped pavement cells of maize leaves these cells are highly elongated and form deep undulations in the anticlinal wall as they mature. The second part analyses the cell shapes of more than 200 species and concludes that strong undulations like those observed in *Arabidopsis* and maize are quite extreme: 95% of the species examined develop more shallow undulations or none at all. This part also investigates the extent to which cell shape features correlate with leaf shape features and the phylogenetic distance. The third part studies more subtle variances in the patterning of the cotyledons of different *Arabidopsis thaliana* ecotypes.

Analysing cell shapes also poses an interesting mathematical challenge: quantifying two- dimensional shapes is a nontrivial mathematical problem as it requires representing the information defined by the coordinates of the outline by only a few quantities. Throughout this project, different metrics have been used alongside principal component analysis to reduce the number of variables. These metrics include traditional morphometric variables, elliptic Fourier descriptors and various other state-of-the-art methods for shape characterisation.

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

It does not exceed the prescribed word limit for the relevant Degree Committee.

Contents

List of Figures

Chapter 1

Introduction

Pavement cells come in a fascinating variety of shapes. Some, like *Arabidopsis thaliana*, are mesmerisingly complex: lobes nonchalantly branching off other lobes, like pieces in a chaotic miniature puzzle game. Neverendingly long maize cells, though neatly ordered parallel to the leaf, hold onto each other with little lobes like teeth on a zipper. Irises have most certainly run out of imagination after painting their flowers: their leaves are paved with long, rectangular cells, like a brick wall. Pretty orchids are oddly boring on the cell level: their leaves are covered in little irregular polygons.

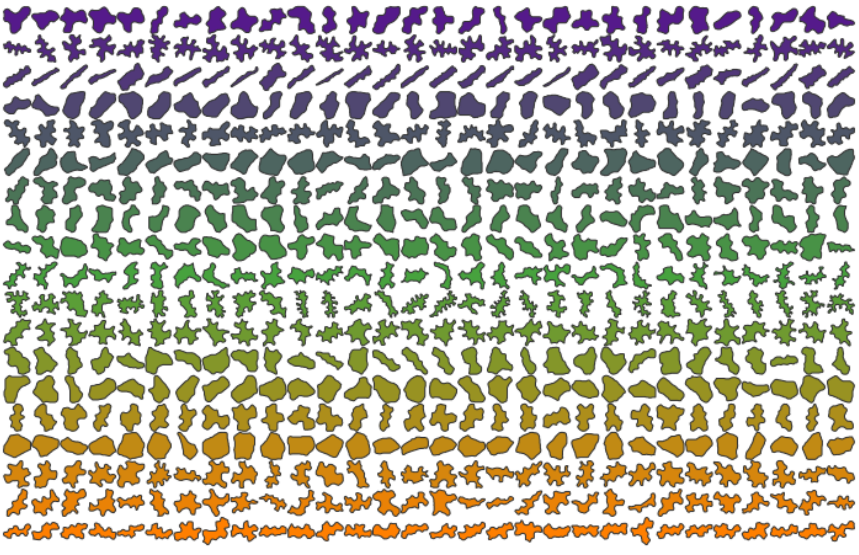


Figure 1.1 – Some examples of cell shapes

The hidden diversity of epidermal cells raises many fundamental ques-

tions about the morphogenesis of pavement cells, including — but in no ways limited to — why and how the shapes might form and how broad the true variety of shapes is.

This project is centred on observing and quantifying the pavement cell shape spectrum. Quantifying shape is an interesting problem on its own: the variety of methods available is nearly as diverse as the pavement cells in the plant kingdom. Chapter 2 introduces the most important ones of these methods and discusses the situations in which they are appropriate to use, in relation to the three datasets presented in this thesis. The first dataset explores the temporal variation of pavement cell shapes in a developing maize plant. The second dataset covers about 200 species from a wide range of vascular plants and explores the diversity on a broad scale. The third dataset examines subtle variations in the pavement cell shape of various *Arabidopsis* lines. Their analyses are presented in Chapters 3, 4 and 5. The rest of this chapter is dedicated to the process that leads to the formation of such complex shapes: the morphogenesis of epidermal pavement cells.

1.1 Living on the edge

The role of the epidermis is to create a boundary between the plant and its surroundings: to protect, to control — but it cannot hide itself. Cells in the epidermis are more exposed than those deeper in the leaf. It is therefore not surprising that even though epidermal patterning is remarkably reproducible on a qualitative level within a species, growth conditions can affect the pattern in more subtle ways: in the cell size or in the depth of lobes.

To provide protection, epidermal pavement cells must grow closely together and restrain each other's growth. At the beginning, immature cells are spherical in shape. The formation of an interlocking pattern requires shape growth in a coordinated fashion. Coordination is the crucial point: as the cells themselves restrict the growth of their neighbours, epidermal patterning really is an emergent property of the whole multicellular system rather than the net effect of single cell growth.

At the boundary, physical conditions change: plant cells are constantly under turgor pressure. In a tissue, the turgor pressure on the cell wall from the cell itself and from its neighbours equilibrates. In the epidermis, while there is turgor pressure from one side, on the other side there is only air. A

change of shape or structure might therefore be necessary to maintain the physical equilibrium.

1.2 Getting in shape

From a purely physical point of view, the shape of a plant cell is the result of a balance between the turgor pressure inside the cell, the tension in the cell wall and the forces due to external constraints at any point along the cell wall. The expansion of a cell is driven by the turgor pressure. Pressure in liquids — such as the cytosol — is homogeneous: this means that the turgor pressure alone cannot be responsible for changes in shape during growth.

The formation of interdigitating cell shapes requires locally varying growth rates, known as heterotropic growth [1] — see Figure 1.2. Internal forces are isotropic: it must then be either the cell wall or the external constraints that are inhomogeneous along the cell wall. Inhomogeneities in the cell wall can be the result of stochastic fluctuations

Before technological advances made it possible to identify the intricate roles biological factors play in forming the outline of cells, most of the efforts to understand how and why cell shapes arise were based on physical principles. The studies discussed below present interesting insight into the early days of cell shape research.

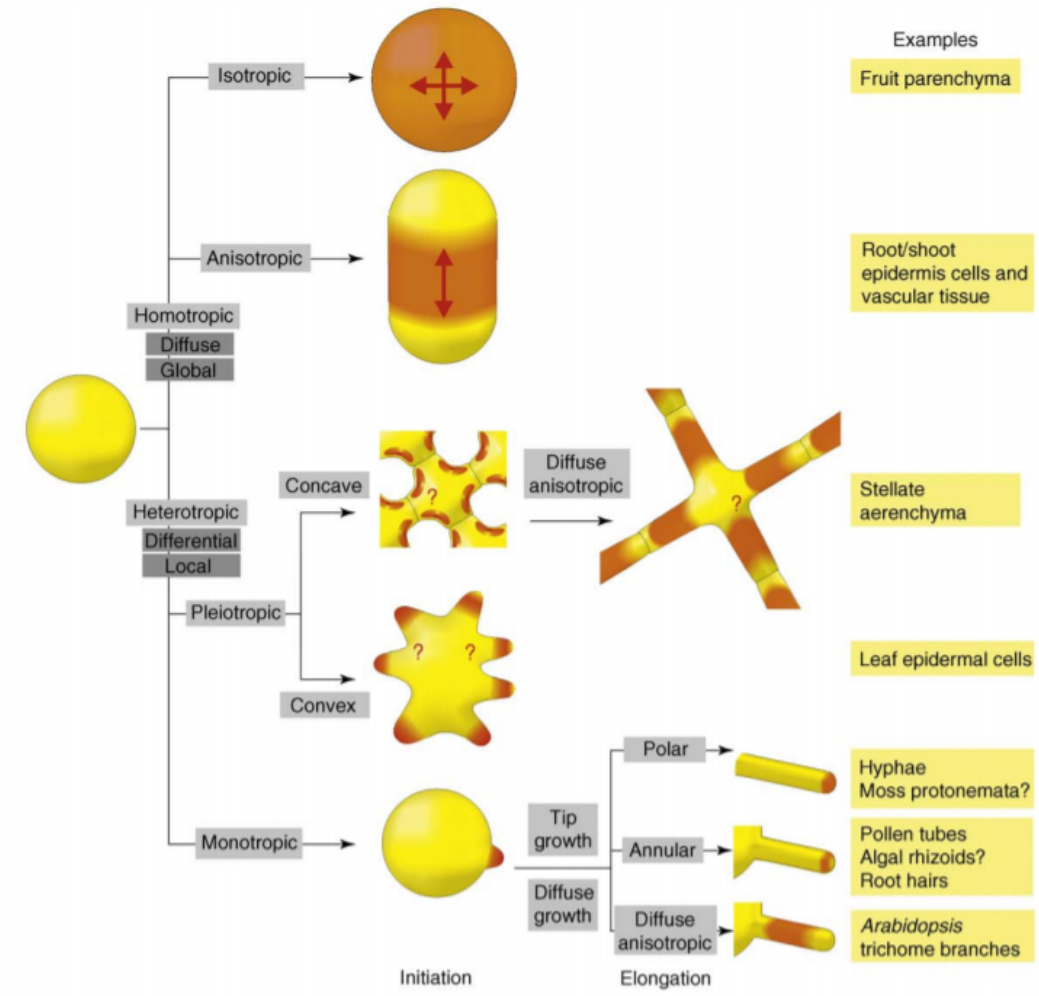


Figure 1.2 – Different modes of growth
Image taken from [1].

1.3 Theories

Counterintuitively, the first cells ever to be observed were not from the outside , but from the inside of the plant: a thin slice of cork, whose honeycomb-like organisation inspired Robert Hooke to coin the term "cell" — still in use 350 years later [2]. The analogy between cells in a tissue and bubbles in foam — compressed spheres — has been the starting point of understanding the three-dimensional shape of plant cells in the tissue for the next three centuries. Unlike actual cells, bubbles are very simple objects whose shape is determined by just two competing forces: one related to the surface ten-

sion and another to the excess pressure inside. The equilibrium shape of the bubble minimises the surface area at a fixed volume. In isolation this is a sphere but when multiple bubbles touch, the situation becomes more complicated. In 1873, from numerous experiments, Plateau deduced four simple laws that describe the angles between the sides of bubbles in contact with each other [3], but a thorough mathematical proof of his empirical laws was only presented a hundred years later [4]. Based on Plateaus laws, Lord Kelvin derived the "ideal shape" of a cell in 1887: a 14-faced body with eight hexagonal and six quadrilateral faces a shape that can fill the three dimensional space while minimising surface area [5]. Lewis wrote several essays on the shape of plant cells in bulk and also on the geometry of ephitelial patterns, finding that in many plants 14-sided plant cells are the most abundant [6]. Marvin compared his observations to those on compressed (originally spherical) lead shots [7]. Matzke compared this result to the number of faces formed in foam made of 2000 individually made, uniformly sized soap bubbles [8]. They found that while Plateau's laws were satisfied in all cases, Kelvin's ideal shape did not appear at all – due to the slight irregularity in arrangement, the equilibrium shapes in the bulk were also more irregular. The average number of faces was, however, very close to the ideal 14 in both cases in the bulk of the material, while on the boundaries the average number of faces was 11.

These simple models were all based on the idea that the boundary between cells should have minimal surface. The presence of a surface has an energy cost associated with it and, by the laws of thermodynamics, every system aims to minimise its total energy. It is easy to see that by reducing the surface area, the system reduces its energy. However, while minimum surface calculations match well the cell shapes in bulk plant tissue, the case of epidermal cells appears to be more complicated as proven by the presence of undulations anticlinal walls. This seemingly unphysical phenomenon was found rather intriguing and the attention shifted from the shape of the bulk cells to those on the periphery.

Early works demonstrate correlations between cell shapes and other features of the plant or its growth conditions: the structure of the mesophyll [9, 10], whether the leaves grow on the sun or in shade [11, 12, 13] or the humidity [14]. Ambronn argues that the lobed boundaries must be the result of locally different surface growth rates [15]. Some authors hypothesise that

the undulations are due to the mechanical conditions during growth. These ideas fall into two different categories: one identifies spatial constraints as the main reason (imposed by the hardened cuticle for example) [16, 17, 18] while the other group considers a dynamic effect: the strain between tissue layers expanding at different rates [10, 19].

1.4 Constraints

The cuticle

Haberlandt explained the undulations of cell outlines as a side effect of growing in confinement [16]. The cuticle consists of an insoluble cuticular membrane, which can contain soluble waxes (cuticular waxes) and can be covered in a layer of soluble waxes (epicuticular waxes). The material that builds up the cuticle is secreted by the epidermal cells. Its softness depends on the composition: unsaturated compounds, when exposed to air, will slowly oxidise and form crosslinks, leading to the hardening of the cuticle. Pfeffer concluded that the cuticle is under tension as epidermal strips from various plants curve outwards when placed in water [17]. Lee and Priestley claim that the hardened cuticle is resistant to tensile forces, thus can restrict the expansion of the underlying tissues [18]. Watson studied the cell shapes and cuticle production in leaves of *Hedera helix* grown under different light conditions [13]. He found that in young leaves the cuticle above the lobes stains differently, but in mature leaves the cuticle appears homogeneous. Based on these findings he hypothesises that the hardening of the cuticle does not happen uniformly and lobe growth happens where the cuticle stays softer. He also assumes that the hardening happens faster when the light intensity is higher: thus sun leaves only have shallow lobes, light shade leaves develop stronger lobes but only at the surface and in dark shade leaves the whole anticlinal wall becomes wavy. The influence of cuticle thickness on cell shape might also be illustrated on *Ranunculus aquatilis*, an amphiphilic plant that exhibits heterophylly [20] — epidermal patterns on aerial leaves (lobed outlines, thick cuticle) and submerged leaves (smooth outlines, very thin cuticle) are different [21, 22].

Avery observed that in tobacco leaves the epidermis drives the expan-

sion in the second half of the growth period and rejects the idea of a non-stretchable cuticle [19]. Korn disagrees with Haberlandt: he argues that if undulations are due to buckling, the observed angles between anticlinal walls should be different and cells should bulge out more on the surface [23]. Higher production of epicuticular waxes was found to correlate with less wavy outlines and more puffed out cells in Arabidopsis mutants [24]. This study hypothesises that the link between the two is the cell wall structure that determines both the cell shape and the permeability of the cell wall for the waxes produced. Similarly, it has been noted that cuticle thickness and epicuticular wax production changes between juvenile and adult leaves in grasses, and this is accompanied by a change in cell morphology [25, 26]. Juvenile leaves produce more epicuticular waxes but the cuticule thickness is only about 1 µm, while adult leaves have a much thicker cuticule, about 3 µm, do not produce epicuticular waxes and cell wall undulations are much deeper [27].

Tissue tension

Different tissues can have different properties: different expansion rates for example. But on the boundary between two different tissues — as in most situations, neighbouring cells cannot really glide on each other — the strain on the tissues must be the same. Whichever tissue is more resistant to stretching will thus limit the expansion and differential tissue growth gives rise to tissue tension. In the context of elongating stems it has been suggested that the internal tissues drive the elongation and the epidermis provides a mechanical constraint based on experiments done on epidermal peels [28, 29, 30].

This idea recently had a renaissance [31], leading to the formulation of the epidermal growth control theory of stem elongation [32, 33]. It was observed that, contrary to expectations[1], in many species the microfibril orientation in the epidermis is not perpendicular to the growth direction [34] – another proposed proof that the elongation of the stem may not be driven by the epidermis, but by different

[1]Composite materials showing uniform fibre orientation are expected to exhibit direction-dependent elastic properties. In particular, the extensibility of the cell wall is expected to be greater in a direction perpendicular to fibre orientation. A simple derivation of this property from basic principles is shown in Appendix A.

tissues. In fact, quite often microfibrils in the epidermis are aligned parallel to the elongation direction [35, 36], which lead to the idea that driving and restricting expansion should be understood along a specific direction: in stems with longitudinal microfibril order, the inner tissues drive the expansion and the epidermis restricts along the length, but radially it is the epidermis that drives the expansion and the inner tissues restrict it [37]. Regarding leaves, it has been suggested that differential growth between the epidermis and the mesophyll – here the epidermis expanding faster than the mesophyll – is responsible for the spongy structure of the mesophyll and the lobed outlines of the epidermal cells in tobacco [19].

1.5 The biochemistry of lobe growth

In the past decades great effort went into exploring the role of the cytoskeleton during morphogenesis and the regulatory mechanisms which enable heterotropic growth. For lobe outgrowth to be possible, the cell needs to be able to restrict growth in the indentation regions while allowing lobes to expand. In recent studies, restriction of growth was associated with cortical microtubules forming bundles around the neck of the lobes [38, 39, 40]. Cellulose synthase complexes are attached to the cortical microtubules in such a way that the orientation of the cellulose fibres produced copies that of the microtubules [41]. It is known that the extensibility of the cell wall is much greater in a direction perpendicular to the aligned cellulose fibrils (see Appendix A) – it seems therefore plausible that the orientation of microtubules directly determines the local growth direction. However, defining the direction itself is not enough: the cell also has to ensure that material is getting delivered into the expanding regions. This is achieved by developing a fine network of actin filaments at the tip of the lobes [42, 43]. The role of the cytoskeleton has been studied mostly in *Arabidopsis thaliana* (reviewed for example in [44]). Lobe formation was found to be suppressed in mutants where the formation of diffuse f-actin structures or microtubule bundles was disrupted. The branching of actin filaments that is necessary for promoting lobe outgrowth is associated with the ARP 2/3 (actin related protein 2/3) complex [45]). Mutations affecting this complex lead to a reduction in the amount and depth of lobes of pavement cells [46, 47, 48, 49] or adhesion de-

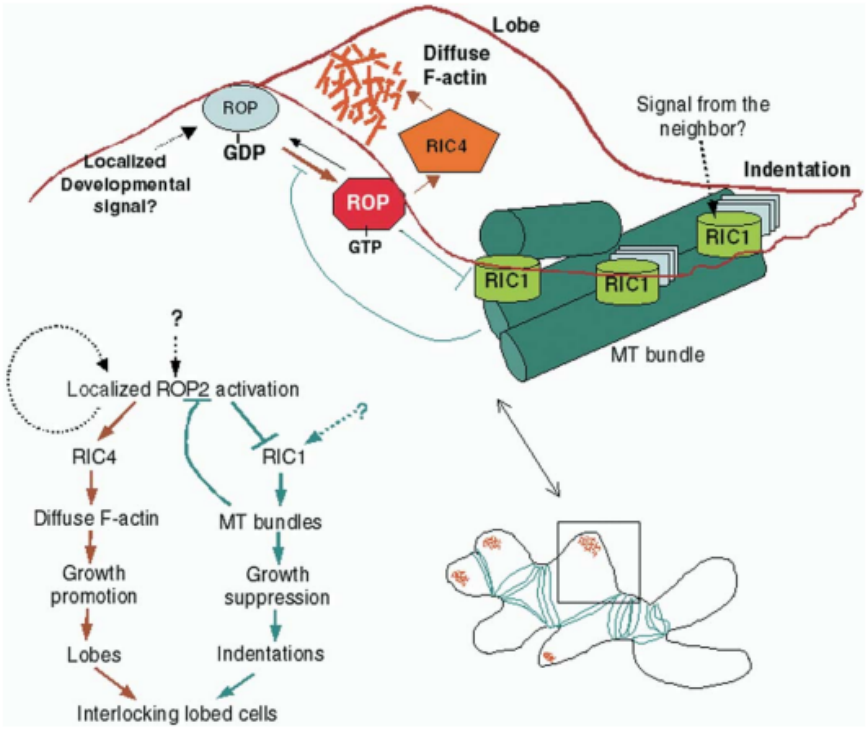


Figure 1.3 – The pathways leading to lobe formation in pavement cells
The ROP6–RIC1 pathway interacts with KATANIN, which induces microtubule reorientation by severing microtubules. Where microtubule bundles form, growth is suppressed — these locations will be the indentation regions. The ROP2-RIC4 pathway promotes the aggregation of fine cortical actin filaments and also inhibits the assembly of microtubule bundles, thus inducing lobe outgrowth. Image taken from [39]

fects [50]. The factor responsible for the activation of the ARP 2/3 complex is the WAVE complex [51, 52]: mutations of the genes encoding its subunits lead to deformations very similar to those where the ARP 2/3 complex is affected [53, 54, 55, 56, 50, 57, 58, 59]. Pavement cells also fail to develop lobes in mutants with microtubule defects [60, 61, 62, 63, 64, 65, 66].

The formation of both the anticlinal microtubule bundles in the indentation regions and the fine actin patches at the tip of the lobes is regulated by members of the ROP (Rho-GTPase of plants) and RIC (ROP-interactive CRIB-motif-containing protein) families in a mutually exclusive fashion [39]. ROP2 and its downstream effector, RIC4, promote the assembly of fine actin filaments and inhibit the ROP6-RIC1 pathway responsible for microtubule ordering [68] — see Figure 1.3. ROP6 and RIC1 interacts with KATANIN

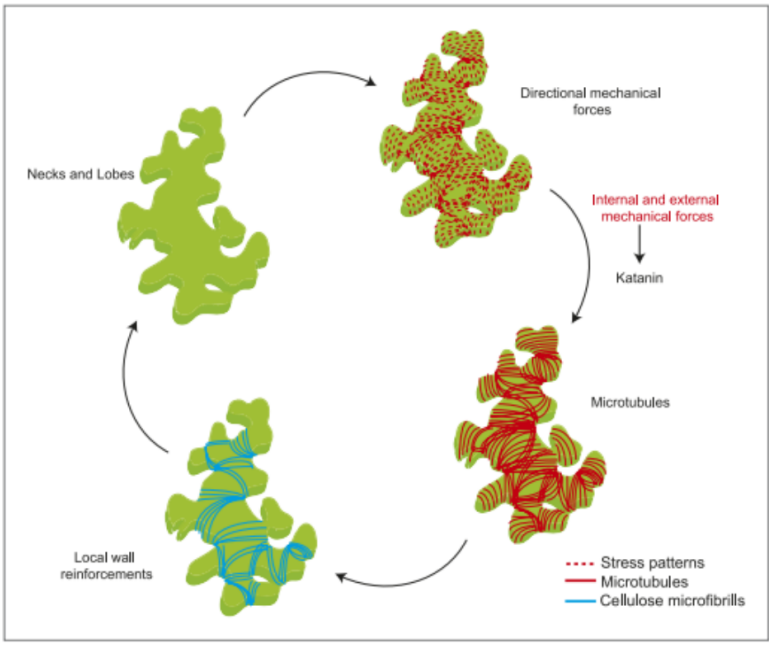


Figure 1.4 – Stress-induced reinforcement
The formation of necks and lobes leads to an uneven stress distribution in the cell wall. Internal and external mechanical forces promote KATANIN and thus lead to a reorientation of microtubules parallel to the lines of highest curvature. The resulting cellulose microfibril orientation follows the microtubules. Image taken from [67].

[66], a protein that can severe microtubules and thus allow reorientation [69]. It was proposed that mechanical stress arising from the curvature of the cell wall also influences KATANIN activity, as a result of which microtubule bundles align with the lines of highest curvature to reinforce the shape of the cell [70, 67]. This is shown in Figure 1.4).

Based on these results, the mechanism driving lobe growth seems straightforward: in some parts of the cell, where ROP6 is activated, anticlinal microtubule bundles create indentation regions and in other parts, where ROP2 is activated, diffuse f-actin ensures that materials necessary for growth get delivered to the tip of the lobes. What complicates the picture is that the indentation regions of one cell must coincide with the lobes of its neighbours and vice versa. In other words, on different sides of the cell wall between neighbouring cells, different ROP proteins must be activated. To explain this polarity, it has been proposed that auxin plays a role in activating the ROPs [71] and that polarity is achieved through the interaction of the ROP2-RIC4 pathway with PIN1 proteins [72].

It is worth noting that the lobes themselves look very different in different species [73, 74]. In some plants (*Arabidopsis*), the whole anticlinal wall is lobed. In others (maize), lobes are only present at the top of the epidermal cells. Characteristic lobe dimensions are also different and so is the base shape of the cells. These differences make it necessary to study the cell shapes of other species too.

1.6 Discussion

Cell shape has been the object of study for centuries as the shape, its formation and function raise mathematical, physical and biological questions, with unsolved problems for modern-day researchers.

The virtue of the early studies is that they consider the shape of the cell in context, and not only at the single cell level. Cell shape studies explored three-dimensional shapes as opposed to surface patterns only. Studies regarding lobed cell boundaries had a strong physical focus, examining the static and dynamic mechanical constraints in a layered structure. Despite the elaborate theories, these papers did not provide sound experimental evidence as it would have been impossible to test them on real plants. Modern studies managed to provide links between the observed cell pattern and the absence or presence of various subcellular elements, but they do no more: to understand the development of cell shapes, we need to be able to track the changing shape through longer periods of time.

The most recent studies make an effort to quantify the observations [40, 75] but only in *Arabidopsis*. For a better understanding of cell shapes, we need to be aware of the differences between species and we need to carry out these examinations in a quantitative manner.

1.7 Conclusions

This thesis is inspired by the variety of epidermal pavement cell shapes observed in plant leaves as well as the breadth of quantitative methods for analysing shape. As there is little data available from species other than *Arabidopsis*, my first goal was to examine other plants and collect a sufficiently large amount of data that, at the start of my work, did not exist. The second goal was to find a method for quantification that is versatile

enough to capture information about datasets showing a large overall variation (Chapter 4) or subtle variations in different ways (Chapters 3, 5).

Chapter 2

Quantifying shape

2.1 Background

Shape is a complex concept that is difficult to define unambiguously. Discussing shape is quite often a matter of semantics: us humans have some sort of a consensus on what we mean by the word "shape", which seems almost too trivial to explain, yet it is not at all straightforward to translate the idea to the language of mathematics. The definition most commonly used in geometric morphometrics — as phrased by Kendall — states that shape is "what is left when the differences which can be attributed to translations, rotations and dilatations have been quotiented out" [76].

The earliest methods for comparing shapes were entirely qualitative: to classify the shape of snowflakes for example, one would consult reference tables [77]. Qualitative classification typically uses similarities that are easy for us to understand but they are also inherently subjective. With computers becoming more and more widespread, the need for unbiased, automatic classification methods arose. However, automatisation requires a quantitative description, which posed a new challenge to researchers: how to say shape in the language of mathematics?

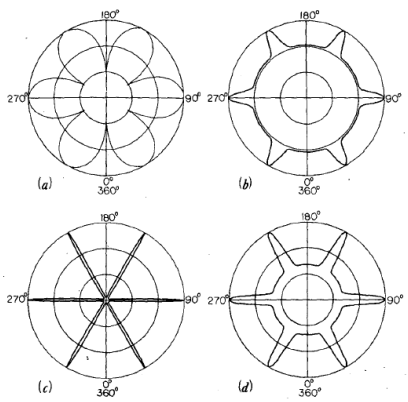
In practice, the shape of an object is defined by its outline. Shapes of different objects can then be compared after normalising the size, eliminating position and orientation. To describe the true shape of the outline, we would ideally want to keep track of all the points along the outline. This kind of shape data is easy to obtain but storing and using it is computationally demanding and, more importantly, it is not straightforward to interpret.

Such a representation requires $2n$ parameters to describe a shape defined by n points in two dimensions — it is unlikely that the data actually contains this many dimensions of information. Moreover, digitisation error would have a very strong influence on the outline. Instead, we can choose from a broad range of methods to reduce the number of parameters needed to represent shapes without losing too much valuable information. In this chapter I present and discuss the most popular methods.

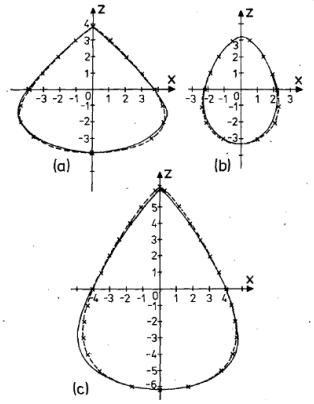
Although cells are three-dimensional objects, cell shapes in the epidermis are usually described in two dimensions, for a number of reasons. Firstly, two-dimensional shapes on the surface of the leaf are reasonably straightforward to study under a microscope. Secondly, the top surface is where interesting shapes appear: leaf cross-sections have, of course, also been examined for several reasons and the cross-section in a plane perpendicular to the surface is usually remarkably uninteresting compared to the lobed outlines of the surface. In often-imaged *Arabidopsis*, the anticlinal cell wall of epidermal cells is simply perpendicular to the surface and the bottom side of the pavement cell is similar in shape to the top side. Thirdly, acquiring three-dimensional cell data requires more complicated imaging and analysis techniques. As such, most of the literature discusses two-dimensional shapes.

2.2 Parametric representation

For simple and symmetric shapes it is often possible to find a parametric function (typically in polar coordinates) that can describe a certain family of shapes just by adjusting a couple of parameters. Snowflakes [78] and raindrops [79] can be described this way with three parameters only — see Figure 2.1. This method has limited applicability but works well in an appropriate case.



(a) Fitting $r = a\sin^2(3\theta)^b + C$ to hexagonal snow crystals. Image taken from [78].



(b) Fitting conical hailstones to $x = \pm a\sqrt{1-(z^2/C^2)}\arccos(z/\lambda C)$. Image taken from [79].

Figure 2.1 – Examples of fitting parametric functions to represent shapes found in nature

2.3 Biological significance: using landmarks

In biological applications, the general shape of the object is usually more complex — not necessarily symmetric or easy to capture by a parametric function. Nevertheless, there are often biologically relevant, distinguishable locations on the shape that can be used to define landmarks. Landmarks allow us to align and scale the shapes, to define descriptors characterising the observed shape variation and to calculate the deformation of the grid that brings one shape to another. An individual landmark is not meaningful in itself: it is the configuration of the full set of landmarks that should capture the shape. Zelditch *et al.* require landmarks to be "(1) homologous anatomical loci that (2) provide adequate coverage of the morphology, and (3) can be found repeatedly and reliably. Two other criteria may also be important under some conditions, that landmarks (4) do not switch positions with each other and (5), in the case of two-dimensional landmarks, lie within the same plane." [80].

Following the changes in the relative locations of the landmarks, corresponding coordinate transformations can be found. These transformations are different at different locations: the net effect can be depicted with the help of warp deformation grids [81]. Figure 2.2 shows the shape of fish species and the warp deformation grid that brings one into another, as pre-

sented by D'Arcy Wentworth Thompson [82] — reproduced image borrowed from [83].

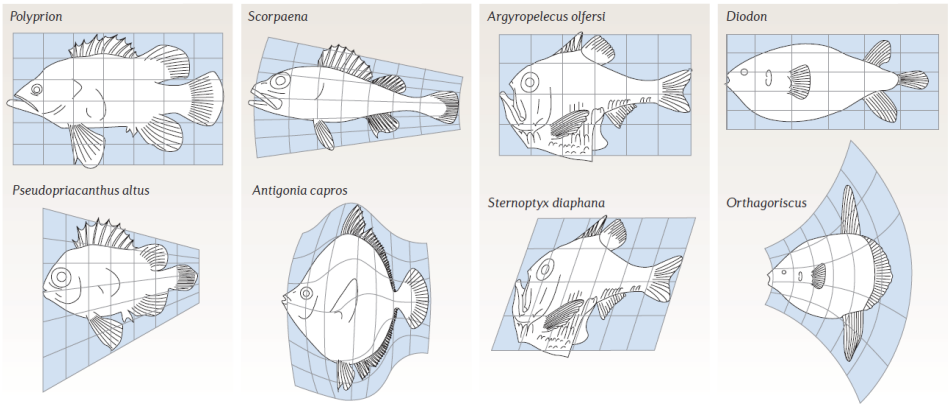


Figure 2.2 – Morphometrics using landmarks - a study on fish species
Original image from [82], reproduction taken from [83]

The applicability of landmarks is, however, limited: for pavement cells, we cannot rely on finding a set of points that can be used to align one pavement cell to another in a manner similar to the various species of fish shown on Figure 2.2 in order to calculate transformation matrices. This method works well in situations where the shape variation is subtle: distinguishing leaves from closely related species [84] or examining the variance within just one [85]. With datasets showing larger variations, the landmarks method becomes suboptimal [86]: landmarks may not be available at a sufficient resolution and as they are usually manually selected, using them on a large dataset can be quite time-consuming.

2.4 Geometric morphometrics

In a more general situation, when landmarks may not be defined, the entire outline of the shape can be used to calculate a number of quantities that are defined to capture certain properties of the shape — with better or worse results. A common feature of these descriptors is that they are defined to be dimensionless, rendering the problem independent of size as required by the mathematical definition of shape.

Because morphometrics evolved simultaneously in several different fields, many of these descriptors exist or had existed under different names. The

same name can also hide different definitions. In the following I summarise the most often used descriptors, along with the definition I adopted.

Aspect ratio

The aspect ratio can be defined either as the ratio of the major axis, a, to the minor axis, b, of the ellipse that fits best to the shape or the ratio of the length and the width of the best-fitting rectangular box. This measure accounts for global characteristics of the shape and ignores local variations. When applied to cell shapes, this measure selects purely based on the elongation of the base shape and does not take the quality of the outline into account. It is also known as the ratio of principal axes [87].

A similar quantity is *eccentricity*: this is also derived from the best fit ellipse and as such, used to measure the elongation of the base shape [88]. In terms of the major and minor axes, the eccentricity is defined as

$$\epsilon = \sqrt{1 - \frac{b^2}{a^2}}.$$

Its value varies between 0 and 1, 0 corresponding to a perfect circle and 1 to a straight line. Some authors, however, define eccentricity as the aspect ratio from above [89].

In my calculations, I define aspect ratio to be the ratio of the minor axis to the major axis, b/a, so the resulting number is bounded within 0 and 1. In contrast with eccentricity, here 1 corresponds to a round object and 0 to an infinitely thin straight line, but the distribution of values is more even in the range epidermal cell shapes realistically fall in. An axis ratio of 1:2 corresponds to 0.5 on this aspect ratio scale, but 0.87 on eccentricity. Given that epidermal cells can, on occasions, develop a 1:10 or even greater axis ratio, using the latter definition of aspect ratio provides a better representation of the variability in the data.

Circularity

Contrary to the aspect ratio, circularity describes the quality of the boundary rather than the base shape. It compares the area and the perimeter of the shape (normalised to a circle)[89]. The (normalised) circularity in this study is defined as:

$$\text{circularity} = \frac{\text{perimeter}^2}{4\pi \times \text{area}}$$

This expression takes a minimal value of 1 for a circle, hence the name. High values of circularity signifies strong local variations of the boundary — e. g. pavement cells with strong undulations in the anticlinal wall. This definition is useful because highly undulating outlines cover a greater range, allowing better distinction between different degrees of undulation. This is also the way circularity is defined in the Momocs package [90] used in the analysis of the datasets.

Variations on the definition include ignoring the normalisation constant 4π; using the ratio of the square root of the area to the perimeter; or inverting the fraction. It is also known under *isoperimetric quotient*, *compactness* [87], *complexity* [40], *form factor* [88], *dissection index* [86] and in palaeobotany, *undulation index* [91].

Ellipticity

Ellipticity is a measure similar to circularity that I constructed in order to better describe the elongated but lobed maize cells. Here the ratio of the squared perimeter and the area is compared against the first harmonic ellipse of the shape instead of a circle. In practice it is calculated as the ratio of the circularity of the shape and the circularity of the first harmonic ellipse.

Similarly to circularity, the value is 1 for a perfectly elliptical shape and increases as the outline becomes more complex.

Solidity

Solidity is another measure that can capture the quality of the outline independently from the base shape. It compares the area of the object to the area of its convex hull, where the convex hull is defined as the minimal convex outline that fits around the original shape.

$$\text{solidity} = \frac{A_{\text{shape}}}{A_{\text{conv. hull}}}$$

This value is bounded between 0 and 1, 1 corresponding to a perfectly convex object and close to 0 we find objects with many deep lobes. In the morphospace of epidermal pavement cells solidity values under 0.4

are rarely observed.

A similar quantity used is *convexity*: this measure is defined as the ratio of the perimeter of the convex hull to the perimeter of the object [87, 89]. Convexity values also cover the 0–1 range and their interpretation is similar to that of solidity. The perimeter is, however, much more affected by pixellisation error and noise in the data. Because solidity does not rely on the perimeter, the influence of these errors is diminished — solidity is therefore more robust against these errors. In medical terminology, *undulation index* refers to a solidity-like quantity (term originally coined by Barrett [92])

Compactness

This measure is another expression attempting to quantify lobes and undulations. It calculates the ratio of the variance in the radius from the centroid of the shape to the area [88]. 0 corresponds to a perfect circle, medium values can mean either a few very deep lobes or several shallow ones. High values correspond to several deep lobes.

The variance of the radius from the centroid can be expressed in other ways too: the *circular variance* computes the mean square error in the radius with respect to the mean radius, or the *elliptic variance* defined similarly for the best fit ellipse [87].

Bending energy

This method first presented by Young *et al.* [93] is based on simple elastic theory (for reference, see [94]). The cell outline is treated as a thin-walled, linear elastic material and the total bending energy is calculated and normalised by the length of the outline:

$$\mathrm{BE} = \frac{1}{L}\int_0^L |K(l)|^2 dl \tag{2.1}$$

where BE is the bending energy per unit length, $K(l)$ is the curvature along the outline and L is the length of the outline (the perimeter).

Similarly to the compactness, the bending energy per unit length also captures local information about the variation of curvature. High values correspond to large frequency and/or amplitude of variations on the boundary, small values correspond to smaller curvatures. This de-

scriptor has a dimension of $(length)^2$, therefore it also depends on the absolute size of the shape.

Interestingly, while bending energy is a physical concept, this measure is far from being unbiological. Bending an elastic material gives rise to a particular stress pattern and plants respond to mechanical stress on a cellular level [67]. Hamant and Moulia suggests that proprioception in plants happens through mechanical signals [95]. If indeed plant cells sense their own shape through mechanical stress patterns, it seems right for researchers to employ a measure that does essentially the same.

Total absolute curvature

Total absolute curvature (TC) captures a similar kind of information as the bending energy per unit length but has the advantage (or disadvantage) of being dimensionless:

$$\mathrm{TC} = \int_0^L |K(l)| dl \tag{2.2}$$

2.5 Series decompositions: one size fits all

While in biological shape characterisation problems methods involving landmarks or simple geometric measures are often preferred as, for example, landmarks themselves can have a biologically significant position; in computerised image processing the use of landmarks is not feasible: the expected variation in shape is very high and datasets are usually too large. In these applications, other, more mathematically complex methods are used to reduce the complexity of the shape to a desired precision.

Moment invariants

Hu introduced seven moment invariants: translation, rotation and scale-invariant functions defined on moments [96]. Khotanzad and Hong argue that recovering the original shape from these moment invariants is computationally demanding, because the original coordinates are not orthogonal , and prefer to use Zernike moments [97]. Zernike moments are even and odd functions defined over a disc of unit radius and form an infinite orthonormal set. They were originally designed to describe optical aberrations in microscope lenses and to

fit this purpose, these moments are defined on a disc of unit radius [98]. Image reconstruction with Zernike moments is very simple: the original image is the weighted sum of the moments. Zernike moments can be used to describe bacterial cell shapes with sufficient precision [99].

Elliptic Fourier Descriptors

Elliptical Fourier descriptors are the two-dimensional equivalent of a Fourier series representation [100]. Any periodic function can be represented as an infinite sum of sines and cosines — its Fourier series:

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos\frac{n\pi x}{L} + \sum_{n=1}^{\infty} b_n \sin\frac{n\pi x}{L}$$

where L is half of the period of the function and $\{a_n, b_n\}$ are the Fourier coefficients. This means that by exploiting the orthogonality of sine and cosine functions, a function (of continuum degrees of freedom) can be represented in a different basis (of continuum degrees of freedom). In practice, we usually only consider the first N harmonics, which still provides a very close approximation of the original function — but possessing only $2N$ degrees of freedom.

Such a transformation can be done in 2D as well. Closed curves can be considered periodic: the period is the total time taken to draw the contour of the shape. Kuhl and Giardina [101] approximates the 2D contour by

$$\begin{aligned} X_N &= A_0 + \sum_{n=1}^{N} a_n \cos\frac{2n\pi x}{T} + \sum_{n=1}^{N} b_n \sin\frac{2n\pi x}{T} \\ Y_N &= C_0 + \sum_{n=1}^{N} c_n \cos\frac{2n\pi y}{T} + \sum_{n=1}^{N} d_n \sin\frac{2n\pi y}{T} \end{aligned}$$

where A_0 and C_0 are the coordinates of the centroid, and a_1, b_1, c_1 and d_1 determine the best-fitting ellipse.

This technique became popular as computer image processing became important and a simple, computationally efficient and omnipotent method was needed [102, 101]. The idea of using Fourier descriptors to quantify shape dates back to 1960 [103]. Over the decades, it

has gained popularity and nowadays this method is widely used across different fields, including plant science.

2.6 Principal Component Analysis (PCA)

With the use of our preferred descriptors we can already reduce the dimensions of our data ($2N$ for a shape consisting of N points) but it is likely that the resulting parameters are more numerous than the information present in the data. Principal component analysis (PCA) is a mathematical tool for reducing the number of variables while conserving most of the variability in the data. This method identifies a set of orthogonal basis vectors that are ordered according to how much they contribute to the variability. The ordering makes it easy to trim the number of variables.

There are different methods to find the principal components: the covariance method is a popular one. The covariance matrix is calculated from the data matrix $\mathbf{M}$ (where rows represent data points and columns represent the original variables) as:

$$\mathbf{C} = \frac{1}{N-1}\mathbf{M}^* \cdot \mathbf{M}.$$

The next step is to diagonalise the covariance matrix by finding the eigenvectors: these will be the principal components. Eigenvalues correspond to the percentage of variability represented. The larger the eigenvalue, the larger is the area covered by the data along that direction. Principal components with small corresponding eigenvalues may be discarded. The final step is projecting the data onto the new set of basis vectors.

On the morphospace defined by for example the first and the second principal components (PC1 and PC2), data points previously overlapping often separate into clusters. This marks the success of finding a "good" representation: using these new parameters we can delimit different classes of data into different regions of the morphospace.

To eliminate biases resulting from the different ranges and magnitudes of variables, it is common practice to centre and scale the original variables. When a variable covers multiple orders of magnitudes, it can be helpful to take its logarithm before centring and scaling.

2.7 Capturing shape and structure

The methods mentioned above were, without exception, designed to compare shape only. Global measures, such as the described geometric quantities, cannot identify where along the outline the differences are. Landmark techniques can only trace the deformation of the shape but are at a loss when new features appear.

As pavement cell shape studies are mostly based on *Arabidopsis*, whose cells develop a very complex lobe system, the need for a method that describes the structure as well arose. Lobe number — often obtained by manually counting — is a key quantity many authors rely on. More recently, lobes are found through skeletonisation, and the lobe number is found by counting the skeleton ends [59].

However, skeletonisation often misses starting lobes and is affected by noise in the outline. Convex hull based methods can overcome this problem. Pauwels *et al.* present a method (originally intended for leaves) where the distance of points on the outline from its convex hull is measured to determine maximum indentation depth. They also calculate the power spectrum of the indentation function, and define a quantity called *lobedness* as the product of the square of the maximum indentation depth and the frequency under which 80% of the total power falls in the power spectrum [104]. Wu *et al.* describe a similar method, where the indentation function is determined, and its peaks are identified as the lobes of the cell [105].

Li *et al.* propose a method for capturing the structure of plants by employing persistent homology [106]. Persistent homology compares topological information about samples. Topology is, by definition, independent of size and focuses on how the points are connected. They apply this method to macroscopic parts of the plant but it is also applicable to cells.

2.8 Potential sources of error

Every step of numerical data processing is a new source of error. In this method, the first source of error is digitally recording the image: this introduces a pixel-scale roughness in the boundaries and an error of ± 1 pixel at the least. When curves are smoothed, we get rid of some of this error, but fine features originally present in the sample may also get lost. Segmenta-

tion introduces a further uncertainty at the boundary: the few pixels next to the actual boundary may be misidentified. The total error in the outline is now around 3–5 pixels, corresponding to a few µm at the magnifications used. As the outline is the main subject of this study, it is important to consider the effect of the error.

2.9 Discussion

It is not always obvious which method is the best suited for the problem at hand. Each method has its peculiarities that may suit certain situations better than others. The available resources (time, computer power) also have to be taken into account as some methods require both.

Whenever it is possible and feasible to define a good set of landmarks, the landmark method is likely to produce good results and helpful visualisation. With landmarks, the shapes can be properly aligned and morphological changes along the whole object can be interpolated. The power of this method is that it builds on biologically relevant features and that it can track local differences in morphology. However, as epidermal pavement cells are generally non-symmetric and have no obvious directionality or any distinguishable features, this method is not applicable in this case.

When landmarks are not available and/or the variation in the data is too large or unknown, series-based methods are a good option as they do not require prior information about the data to function well. It is important to keep in mind that Zernike features are defined on a circle: outlines of heavily elongated shapes will carry less weight in this description. Fourier descriptors are very widespread and computationally efficient, with many different implementations available for free.

An advantage of series-based methods is that shapes can be reconstructed from the component values: this allows not only a comparison between the original shape and the one reconstructed from the final variables, but shapes corresponding to any point in the morphospace can be calculated. However, non-periodic undulations are often suppressed as their contributions spread out over several harmonics. This may pose a problem with the analysis of pavement cell shapes.

In a situation where no natural landmarks are available but the quality of morphometric variations is known to some extent, geometric descriptors

can provide a simpler and more efficient representation, at the loss of reconstructibility. Peura and Iivarinen propose the use of simple morphometric descriptors as opposed to Fourier descriptors unless the shapes are truly periodic [87]. They argue that the measures used should be simple, concise and easy to interpret for a human being: a mathematical phrasing of qualitative descriptors.

In light of the expected variation in cell shape, I tried a set of geometric descriptors as well as elliptic Fourier descriptors on a set of example shapes resembling epidermal pavement cells.

2.9.1 Comparing geometric descriptors

As it was straightforward to choose aspect ratio for quantifying elongation, I wanted to know which of the quantities described above can capture variations in the outline independently of elongation. Figure 2.3 shows the values of the geometric descriptors for a set of example shapes drawn to resemble those observed in the epidermis. Panels **a**–**f** show the example cells coloured according to the circularity, ellipticity, solidity, compactness, bending energy per unit length and total curvature respectively. On each panel, the first row contains slightly rectangular shapes of varying elongation (aspect ratio). The second and third rows show shapes with shallow and deeper "lobes" and the fourth row shows example shapes with a slightly more complicated branching structure. Along the horizontal axis the aspect ratio takes the approximate values of 1, 0.8, 0.6 and 0.4.

Circularity (panel **a**), true to its name, captures the deviation from an ideal circular shape a little to well: the noticeable colour change in the horizontal direction indicates that circularity varies strongly with aspect ratio. The rightmost shape in the first row and the leftmost shape in the second row have similar circularities but very different shapes: the first one has no undulations but it is highly elongated, while the other shows mild undulations and no elongation.

Ellipticity (panel **b**) efficiently corrects this characteristic: we see almost no variation horizontally. Solidity (panel **c**) is also independent of aspect ratio. Both of these measures clearly distinguish the different types of outlines, but in a different manner. Ellipticity increases from shallow undulations through deep undulations towards complex lobe. In contrast, solidity orders rows 2–4 differently: complex lobes are placed between shallow and

deep undulations.

Compactness (panel **d**) orders undulation depth/complexity similar to solidity but also varies very strongly with elongation. Bending energy per unit length (panel **e**) also varies with aspect ratio, and this variation becomes stronger as the complexity of the outline increases. Total curvature (panel **f**) is a lot less dependent on the aspect ratio than bending energy.

Based on this, ellipticity and total absolute curvature were the quantities that captured the complexity of the outline according to my personal qualitative judgment.

With respect to error, it must be noted that quantities derived from the perimeter are more vulnerable to pixellation error than quantities based on area. Compactness is derived from the variance in the length of the radius. The absolute error in the radius and in the deviation from the average radius are the same, but the fractional error is much greater in the latter, making compactness even more influenced by the error in the outline. Any additional roughness in the perimeter is amplified in its derivatives. Curvature is defined in terms of the first and second derivatives, rendering total curvature and bending energy the least robust against error in the outline. For this reason, I decided against using compactness or the curvature-based quantities.

In terms of practicality, circularity and solidity are built-in functions in the Momocs R package [90] which was already used for calculating elliptic Fourier descriptors. I defined functions that calculate ellipticity, compactness, bending energy per unit length and total curvature, however these involve more calculations and took longer to run. This extra time was noticeable with the vascular plant dataset containing more than 10000 cells.

As solidity measures undulation depth independent of the aspect ratio and circularity measures the complexity of the outline, I finally decided to use both of these measures alongside aspect ratio and area as a default, and perform principal component analysis to determine the relative importance of these quantities in describing the dataset.

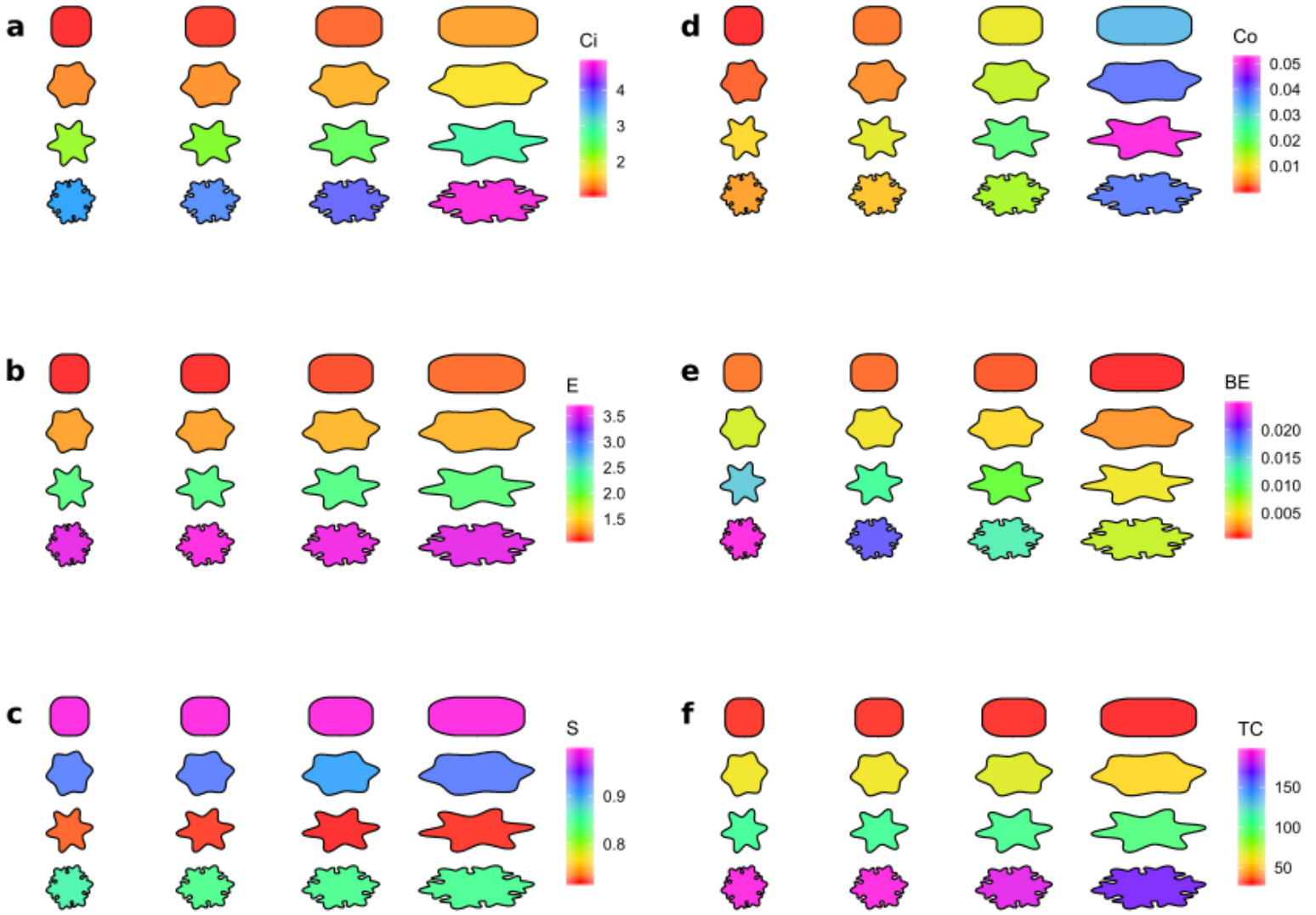


Figure 2.3 – **Testing geometric descriptors** — A set of example cells coloured by **a** circularity (Ci), **b** ellipticity (E), **c** solidity (S), **d** compactness (Co), **e** bending energy per unit length (BE) and **f** total curvature (TC). The aspect ratio of shapes in columns 1–4 is 1, 0.8, 0.6 and 0.4 from left to right.

2.9.2 Testing elliptic Fourier descriptors

Elliptic Fourier descriptors were calculated for the same set of example shapes up to the first 12 harmonics. Principal component analysis was performed on these descriptors and the shape variations corresponding to the first 4 principal components is shown in Figure 2.4.

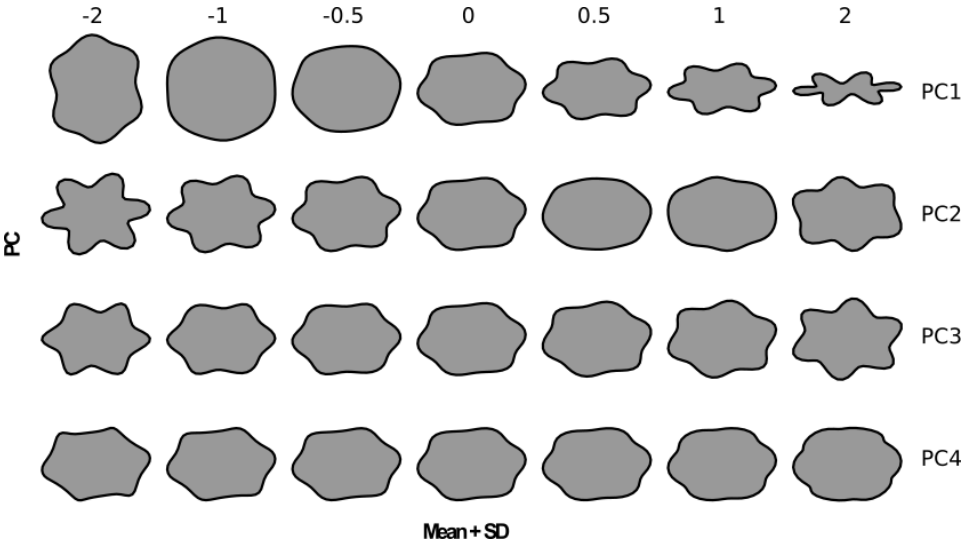


Figure 2.4 – Testing Fourier descriptors — Shape variations corresponding to the first four principal components derived from the elliptic Fourier descriptors calculated for the example shapes shown on Figure 2.3.

All four principal components seem to capture the six-fold symmetry of the example shapes. PC1 also captures the elongation. PC2, PC3 and PC4 are more difficult to interpret, but the reconstructions showed a promising resemblance to the original cells.

It must be noted, however, that the complexity of cell shapes in the fourth row appears to be lost. This phenomenon is further discussed in Chapter 6.

2.10 Conclusions

Epidermal pavement cells in general are quite diverse in shape. As there are two main features that describes a general dataset: elongation and undulations in the outline, my aim was always to find two descriptors on the basis of which the shape of cells can be quantified and, in a successful case, classified.

Since landmark methods are not appropriate for this problem and pavement cells in general lack symmetry, for the problems presented in the following chapters I opted for geometric descriptors and also tested Fourier analysis. Both of these methods, when necessary, were coupled with PCA, in which case the first two principal components were chosen as the two variables.

Methods offering structural information were not found necessary in the case of studying maize cells as maize cells do not develop a complex branching structure, only reasonably periodic undulations (Chapter 3); or in the case of comparing cell shapes from several plant species, where differences in the global characteristics are already quite diverse (Chapter 4). In these case, the information gained from structural data is little compared to the increased complexity of computations.

Chapter 3

Morphogenesis in maize

3.1 Motivation

Maize (*Zea mays*) epidermal cells — shown on Figure 3.1 — are particularly interesting: they are unusually heavily elongated and display surprisingly periodic undulations. These undulations, however, only appear on the longitudinal sides of the cell and close to the surface, unlike in *Arabidopsis*, where undulations reach the full depth of the anticlinal wall.

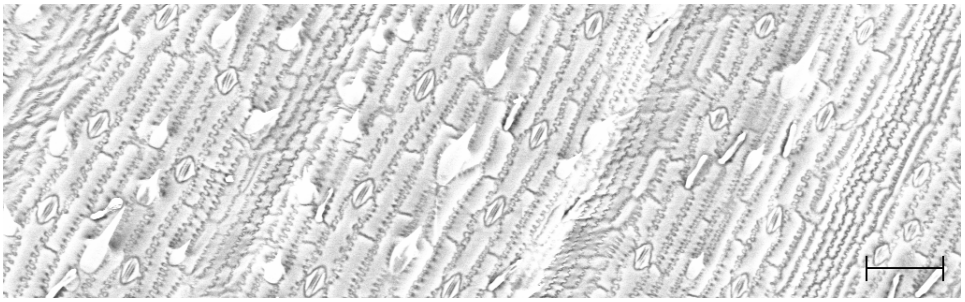
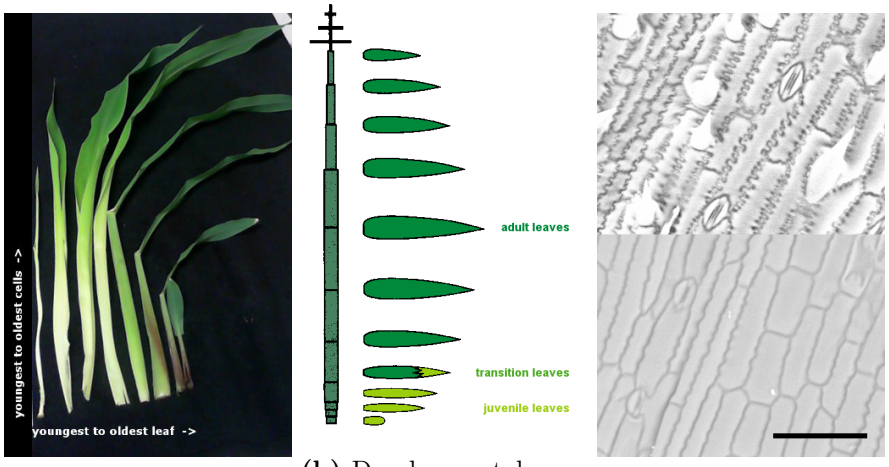


Figure 3.1 – Complex cell shapes in maize leaf epidermis.
Scanning Electron Micrograph, scale bar = 100 µm

Young cells tend to have a shape as close to spherical as their neighbours allow. Shape formation therefore relies on the ability of plant cells to grow non-isotropically. The leaves of maize are, like in many monocot species, linear in structure. The growth zone at the bottom of every leaf is also organised linearly, allowing us to follow the stages of maturation from the base of the leaf (least mature) to the tip (most mature) in a sequential manner — see Figure 3.2a. In mathematical terms, the distance along the leaf can

be used to represent the degree of maturation, making maize a convenient choice for studying morphogenesis — in dicot species the correspondence between location and degree of maturation is not so straightforward, [107].

Cells at the very bottom of the leaf form the division zone: here the cells elongate a little and divide to form new cells of similar proportions. Further along the leaf, cells transition into the expansion zone, where they keep elongating without dividing anymore, until they reach the maturation zone, where elongation stops. The boundaries of these zones vary in time as leaves eventually reach a final size — at this point, cells do not elongate anymore. In a growing leaf the typical extent of the division zone is 1–2 cm and the width of the expansion zone is around 5 cm [108].



(a) Age and location **(b)** Developmental phases **(c)** Mature cell shapes

Figure 3.2 – The development of maize leaves
(a) Separated leaves of a three-week-old maize seedling: the oldest leaf is the outermost leaf, the youngest is the innermost. Within a leaf, young cells are located at the bottom, mature cells at the top. **(b)** Location of leaves in juvenile, transition and adult phase in a maize plant. Schematic drawing adopted from [26]. **(c)** Mature cell shapes in juvenile (bottom) and adult (top) leaves. SEM images taken at 400x magnification, scalebar corresponds to 100 µm.

Mature cell shapes can, however, differ from leaf to leaf. In the development of plants we distinguish three stages: a juvenile vegetative, an adult vegetative and reproductive phase [109]. The location of the vegetative phases is shown on Figure 3.2b. Typically in wild type maize, the first 4–5 leaves are in the juvenile phase and the next 2–4 are in a transitional phase. Differences between the vegetative phases can be subtle in some species, but in others — like maize — juvenile and adult leaves show

different traits: these species exhibit heteroblasty [110]. Differences are observed in leaf structure and leaf shape [111]; cuticle thickness, epicuticular wax coverage and hair formation [27]; and most importantly, epidermal cell shapes [26, 112, 25]. Figure 3.2c shows SEM images of the mature adult and juvenile epidermis.

While there are studies available that quantify variations along the developing maize leaf in epidermal cell length [113, 108], quantitative studies of the epidermal cell shape (as viewed from the top) are scarce — authors typically present a set of images displaying the differences in shape between immature and mature cells in the juvenile and the adult phase, and qualitatively discuss that the main difference between the shapes of mature juvenile and adult cells is the depth of crenulations in the anticlinal cell wall [112, 114, 27]. Bongard-Pierce *et al.* measured the variation in the cross-sectional shape of epidermal cells and found that juvenile ones are more circular in cross-section, mature ones are more tightly packed and thus have a rectangular cross-section. They quantified this difference using a measure equivalent to circularity — see Chapter 2 for definition.

Quantitative studies of lobe formation in the anticlinal cell wall have been carried out in *Arabidopsis* [40, 75] but not in maize. In maize, the relatively short length of the division and expansion zones make studying the cell morphogenesis in real time hardly possible. At the base, leaves are tightly wound around each other, hiding the growing cells from the curious eyes and conventional imaging methods. However, the presence of the developmental gradient along the leaf allows us to convert time into distance, simplifying data collection.

3.2 Aims

Since there are no studies characterising the development of the interlocking cell outline in maize, the primary aim of this project was to collect cell shape data from maize plants at different ages, from different leaves and at different locations on the leaf and develop a numerical method for quantifying cell shape that captures the variations appropriately.

Leaves from seedlings aged 1–4 weeks were sampled (juvenile and transition phase leaves) and cell shape characteristics (qualitative and quantitative) were recorded along with the location on the leaf, the number of the

leaf and the age of the seedling.

Section 3.3 presents the experimental methods used, 3.4 summarises qualitative observations. Section 3.5 explores a quantitative method using this dataset and 3.6 presents an extended qualitative analysis, based on similar principles. The data used is supplied in Appendix B.

3.3 Experimental methods

3.3.1 Plant material

Maize (wild-type 'B73') plants were grown in Levington M3 compost with 16 h light/8 h dark cycles in the greenhouse of the Sainsbury Laboratory. Day and night temperatures were 26 °C and 19 °C respectively. During 'daytime', light intensity was regulated between 125–650 $\mathrm{W\,m^{-2}}$.

3.3.2 Imaging

Images were collected in a Zeiss EVO HD15 Environmental Scanning Electron Microscope (SEM) at 20 kV, under high vacuum. Samples of approximately 1 cm × 1 cm were cut from fresh leaves, fixed on a sample holder by carbon tape with the adaxial side up and kept at −25 °C during imaging using a Peltier cooler to better preserve the cell shapes.

Imaging modes The microscope is equipped with a backscattered electron (BSE) detector and a secondary electron (SE) detector as well. Backscattered electrons are electrons that do not enter the sample because they suffer reflection at the surface, typically in a direction normal to the sample surface. The BSE detector is located right above the sample for this reason and produces a 2D-looking image as a result. Secondary electrons are produced by the interaction of the incoming electron beam and the sample, usually scattered of at a larger angle. The SE detector is therefore located on one of the sides and produces a much more topographic view. In all cases, both images were captured and saved.

Magnification Using the same magnification makes it easy to compare features. It was found that at 400× magnification we see enough detail of the individual cell shapes but also capture a number of cells at the same

time. Images for analysis were mostly taken at this magnification. Other images were taken at smaller magnifications to show a larger segment of the sample and at higher magnifications — usually of an interesting feature.

Resolution Initially, images were taken at a resolution of 0.72 µm corresponding to dividing the target area into 1024×768 pixels at $400\times$ magnification. It was later found that this is not quite sufficient to capture the very fine shape of the lobes of mature maize pavement cells at $400\times$ magnification. As using a larger magnification would have meant capturing very few cells in an image (or even risk not being able to image a single entire cell), two experiments were carried out to determine whether increasing the resolution helps. In these experiments the resolution was increased two– and three-fold for imaging the same target area at the same magnification. Unfortunately, scanning at higher resolution made the system unstable, especially when using dual imaging mode. While the visibility of fine lobes did improve with higher resolution, dust grains and ice crystals on the surface also became more visible, diminishing the benefit of high resolution imaging.

3.3.3 Segmentation

The first step of the analysis was to obtain the cell wall network from the images taken. It was obvious from the beginning that the SEM images were too noisy for a simple thresholding. Intelligent curve tracing algorithms in ImageJ [115] and GIMP ([116]) were tested with little success. CellSet ([117]), a software developed for the segmentation of confocal images gave good results for younger cells, where the lobes were still shallow — for this purpose, BSE images were inverted (as CellSet is looking for a bright boundary on black background). This method failed for more mature cells because it tries to minimise the curvature of the cell wall outlines, and while doing so, it often cut through the necks of deeper lobes. CellProfiler ([88]) also failed to identify mature maize pavement cells at a wide range of settings tried. It was found that tracing the boundaries by hand was the quickest and most accurate way to extract the cell wall network from the SEM images.

Hand segmentation was carried out on a Microsoft Surface 2 tablet using GIMP ([116]). The stylus originally provided with the tablet was used as an input device. After loading the image, a new transparent layer was created above the image layer and the cell outlines were traced using the freehand

drawing tool with the linewidth set to two pixels, hardness set to 100 and brush dynamics turned off to obtain a black-and-white image. Once all cell walls were traced, layer containing the cell wall network was placed above a white layer and the image was saved in bitmap format. Cell wall outlines were then extracted from the binarised bitmap images in MATLAB [118] using the `bwboundaries` function and stored as a list of 2D coordinates.

3.4 First glance

Figure 3.3 shows a representative selection of SEM images from seedlings aged one, two, three and four weeks. Images shown here are BSE images, lacking topographical information. Contrast and brightness varies a lot between images, mostly due to ice crystals, dust and bursting cells (dark ones). Despite the suboptimal quality of the images, the outline of the cell wall network is visible enough to infer some trends:

Within the leaf

Leaves were examined at several spots between the ligule and the tip of the leaf. Close to the ligule, the cells are square-shaped, their boundaries are smooth — this is where the young cells are located. Towards the tip, cells first elongate, while boundaries stay smooth. Lobe growth first appears in the cells located between veins. Later the cells above the veins start interdigitating as well. As cells mature further, lobes become more pronounced. The maturation of cells happens over a typical distance of 5 cm, this seems to be independent of the age of the plant and the location of the leaf.

From leaf to leaf

Average cell size seems to vary between leaves of the same plant. Older leaves tend to have bigger cells, younger leaves have smaller cells. On the other hand, younger leaves seem to exhibit deeper lobes than older leaves — this is because younger leaves on 3– and 4-week-old plants are in a transition phase and below the very top part the mature cell shape is the more strongly interdigitating adult shape, while the older leaves are juvenile leaves, whose mature cell shape is less crenulated. See Appendix B for examples.

With age

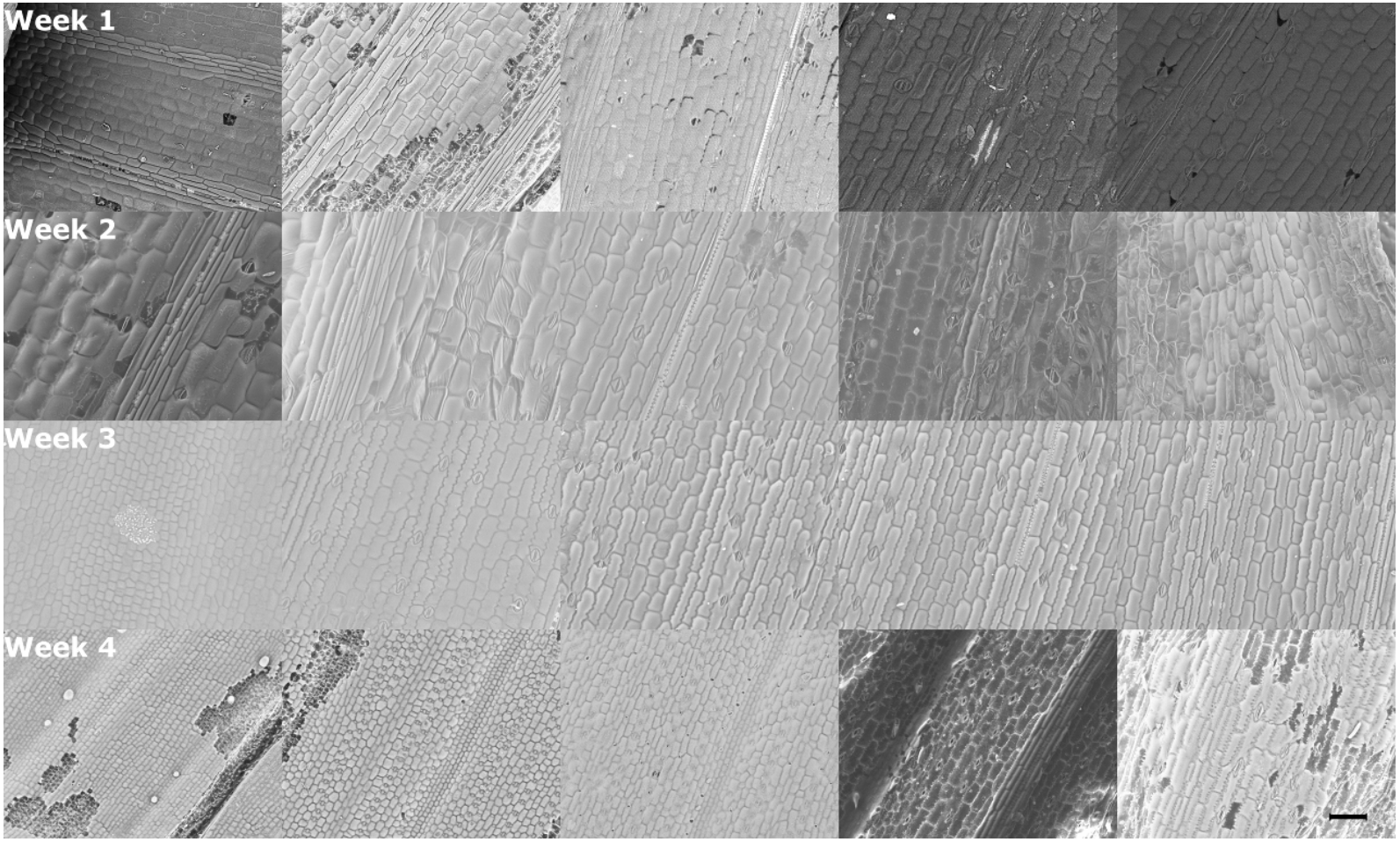


Figure 3.3 – **Variation of maize epidermal cell shape with age** — rows show SEM images from one–, two–, three– and four-week-old seedlings respectively. Images are ordered within a row according to the location on the leaf: the leftmost image being closest to the ligule, the rightmost image closest to the tip of the leaf.
The scale bar is 100 µm and applies to the two–, three– and four-week-old samples, where images were taken at 400× magnification. Magnification of images from the first week are: 241×, 244×, 269×, 417×, 376×; from left to right.

Cell outlines develop slight undulations in one-week-old seedlings already, but distinguishable and more periodic lobe growth appear at about three weeks, on leaves transitioning to adult phase. It takes four weeks until the first mature adult cells appear, after this time cell shapes do not change significantly. See Figure 3.3 for the schematics of maize epidermal cell shape formation based on these observations.

3.5 Quantitative analysis: proof of principle

Following the qualitative observation that cell shapes in maize differ in size, elongation and undulation depth (see Figure 3.4), my aim was to quantify these properties using some of the measures described in Chapter 2. Size can easily be expressed as the area of the cell and the elongation as the aspect ratio but undulation depth does not naturally correspond to a measure. Instead of finding one measure for quantifying undulations, I calculated circularity, compactness and variation of curvature as defined in Chapter 2.

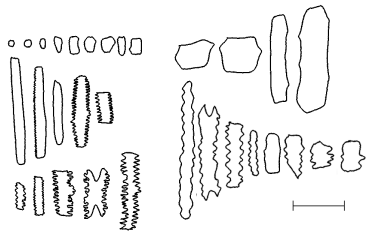
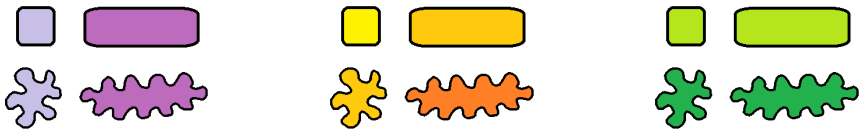


Figure 3.4 – The range of shapes and sizes observed in maize epidermal pavement cells. Cells selected to represent the variety of cell shapes seen during the development of the maize plant, taken from leaves 1-7 of a three-week-old and leaves 7-9 of a four-week-old seedling.
Scale bar: 100 µm

All three of these measures are defined with respect to a centre, meaning that they work best for quasi-circular objects. The elongation of maize cells shows a tenfold variation, making it hard to distinguish whether the value of a certain variable is due to the undulations or the elongation. To overcome this problem, I defined ellipticity similar to circularity, as defined in Chapter 2. As Figure 3.5 shows, this measure is better at distinguishing undulating and smooth boundaries independent of elongation.

Figure 3.6a shows the outlines from six different groups of cells taken from the transition phase leaves of three– or four-week-old wild-type maize



(a) Aspect ratio **(b)** Circularity **(c)** Ellipticity

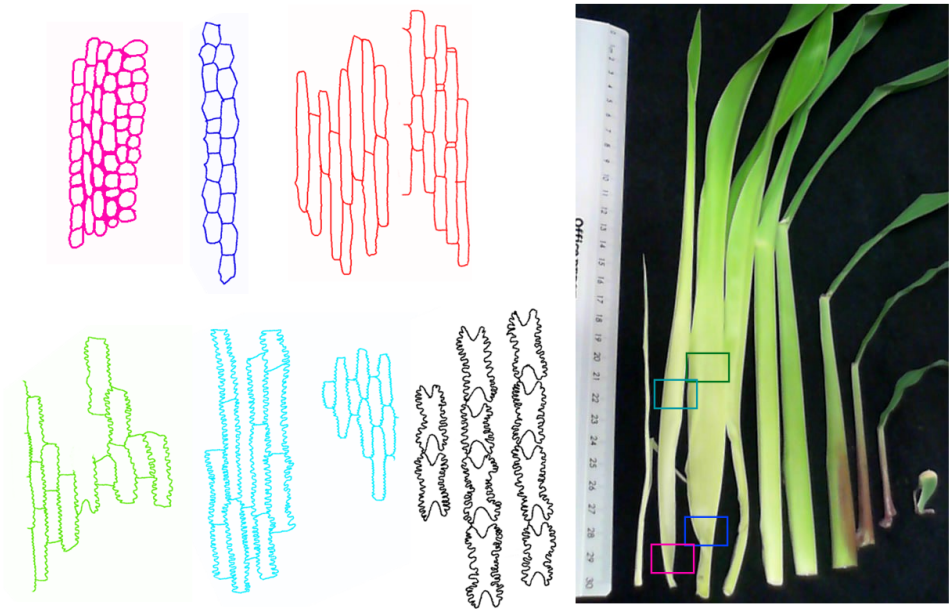
Figure 3.5 – Four shapes, three measures — same colour marks similar values. **(a)** Aspect ratio ignores the boundary and separates based on the overall cell shape. **(b)** Circularity separates wiggly and smooth outlines provided the base shape is similar but it is influenced by the aspect ratio. **(c)** Ellipticity separates based on the quality of the boundary and is not influenced by the overall shape as much as circularity.

plants in different colours. The photo shows the leaves of a 4-week-old maize seedling that was sampled. The blue and pink boxes show a 1 cm segment where the cells showed in pink and blue originate from. Cells in red were also take from the blue segment — these cells are located above the veins of the leaf. Cells in green and turquoise were taken from between 7 and 8 cm from the ligule. Cells in black are interstomatal cells from leaf 7 of a three-week-old seedling, at approximately 11 cm. These areas were selected to represent the full range of cell shapes observed and cell shapes were grouped into the six groups depicted on Figure 3.6a based on a visual classification. The aim of this preliminary analysis was to see whether qualitatively different shapes can be separated on a purely mathematical basis.

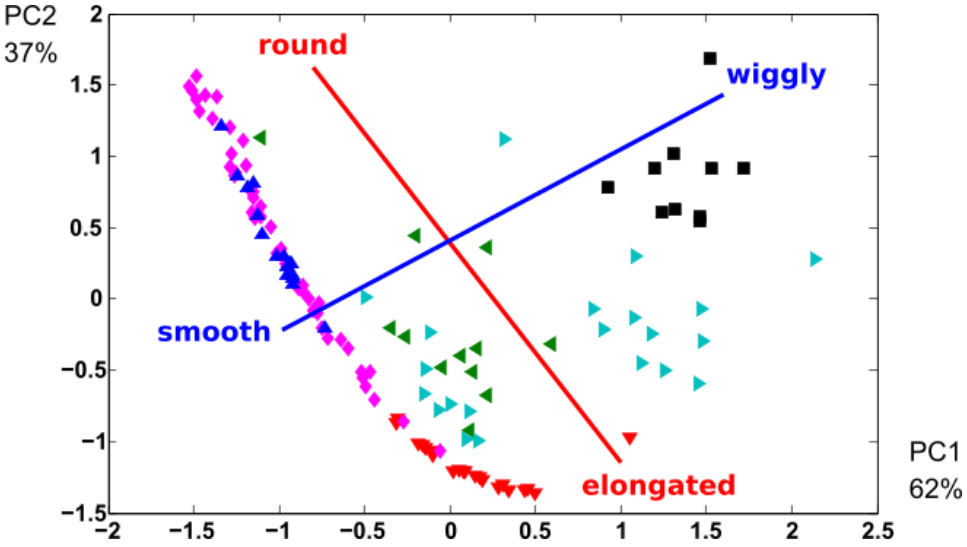
Morphometric descriptors were calculated in MATLAB [118] from the extracted outline coordinates. The morphometric descriptors used in this preliminary analysis were aspect ratio, circularity, ellipticity and compactness as well as absolute area. In order to reduce the number of variables, the resulting dataset was subjected to principal component analysis (see Section 2.6).

Results are shown in Figure 3.6b. The data is well represented by two principal components: the first principal component (PC1) accounts for 62%, the second principal component (PC2) accounts for 37% of the variability of the data. Figure 3.6b locates every cell on the basis of the two principal components. The colour of the data points corresponds to the grouping.

The success of this method is marked by the clustering of data points from the same dataset: similar outlines are represented by similar values of PC1 and PC2. Moreover, two trends can be observed: cells are separated based on their elongation along the red line: round cells are situated closer



(a) Six types of cell shapes used for analysis
Outlines obtained from SEM images of three– and four-week-old wild-type maize seedlings, grouped by average shape. Coloured boxes indicate the approximate location of the cells on the plant.



(b) Cells from Figure 3.6a characterised by the two principal components

Figure 3.6 – Preliminary shape analysis — qualitative descriptors can be expressed quantitatively. Round shapes separate from elongated ones along the red line, smooth outlines separate from wiggly along the purple line. Cells from the same groups form clusters.

to the top, elongated cells closer to the bottom. Roughly perpendicular to this, the purple line shows the variation of boundary quality: smooth cells sit towards the left, wiggly cells — especially interstomatal cells — can be found on the right.

In this part of the project it was shown that the method using traditional morphometric descriptors coupled with PCA was able to define two parameters that allow a distinction between qualitatively different cell shapes on a quantitative basis, without having to choose from the broad range of descriptors available. This routine is reproducible and can give good results when the shapes in the data set differ in multiple ways.

3.6 Along and across leaves

In order to analyse the variation in cell shapes along and across leaves in seedlings of different age, twenty cells were segmented from every sampled location. The number of the cells used was chosen to be as high as possible, taking into account the time-consuming nature of manual segmentation(see Section 3.3.3) and the number of cells per image at the optimal magnification. A few samples yielded poor quality images or fewer than twenty intact cells and were thus not included in this dataset. The final set contains 1620 cells from 81 locations. The list of samples, along with the outlines of the cells used in the analysis can be found in Appendix B.

3.6.1 Methods

In the highly linear maize leaf — especially towards the midrib — cells are also highly linear in structure. This means that the base shape of the cells is uniformly rectangular, so the variations only include the elongation of the cell and the undulation depth of the longitudinal cell wall. Moreover, as the growth is highly anisotropic, the width of the cells, w, does not change much compared to the length l.

Assuming that the undulations can be described by a periodic and symmetric function, which is supported by observations, we can easily calculate the approximate solidity of the cell in terms of the maximum cell width and the undulation depth d. As shown on Figure 3.7, we can divide the cell into three regions. Due to the symmetry of the outline, the outer areas occupy an area of approximately $0.5dl$ each, while the inner part is simply a rectangle

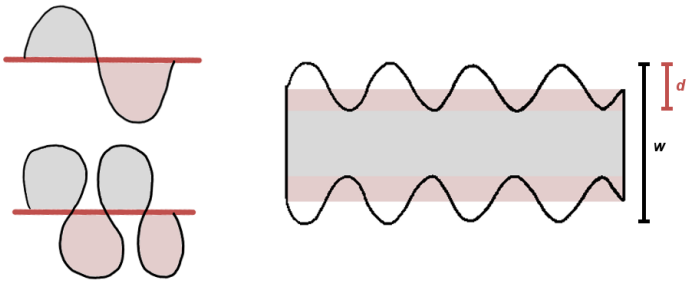


Figure 3.7 – Parametrising the maize cell — The outline of maize cells on the longitudinal side can be approximated with a symmetric and periodic wave. **Left** Examples of waves: the grey and the pink areas are equal for a symmetric wave. **Right** Pink rectangles are equal to the area from lobe minimum to lobe maximum due to symmetry. The sum of the grey and pink areas is the area of the maize cell.

of area $(w - 2d)l$. Approximating the convex hull area of this shape as wl, the solidity of a cell with undulation depth d is $1 - d/w$. Elongation can be quantified with the aspect ratio, defined as w/l.

Cell area was also measured (in µm^2)to follow the variation of absolute size. Samples were grouped by the age of the seedling, the number of the leaf and the location on the leaf. Location was expressed as the distance from the ligule (in cm). Data processing and analysis was carried out in R, using the ggplot2 and Momocs packages in addition to the standard build [119, 120, 90].

3.6.2 Elongation

It is expected that from the base of the leaf the elongation of the cells increases and thus the aspect ratio decreases. For most samples, the data follows the expected trend: aspect ratio starts from high values corresponding to round/square-shaped cells and steeply drop over the course of 5–10 cm. In some cases the first data point has a lower value than the second: this is explained by the transversal variation in elongation across the leaf. Cells situated above or close to veins are often much more elongated than the cells between veins. In general, care was taken to choose areas roughly at the same transverse location with respect to veins but on occasions this was unavoidable. Overall, the variation of the elongation matches previous qualitative observations [121].

3.6.3 Undulation depth

Solidity close to the base of the leaf is expected to be high as the cell outline here is usually smooth, no undulations are present. In juvenile leaves, solidity is expected to moderately decrease further away from the leaf base (moderate undulations in the anticlinal wall), while in adult leaves we expect to see a more pronounced decrease as the undulations of the mature cells are much deeper.

It is apparent that the drop in solidity happens over a longer distance than the drop in elongation: the increase in undulation depth is part of the maturation process. The difference between maximum undulation depth on different leaves matches expectations. Cells on leaf 1 show only very weak undulations: minimum average solidity is just below 0.9. On leaves 2 and 3, minimum average solidity reaches 0.85 and on leaves 4–6 drops as low as 0.8, corresponding to a slowly increasing undulation depth. In leaves 7–9, minimum average solidity reaches 0.65, indicating a sudden increase in undulation depth, which points at the adult-like characteristics of mature cells expected in transitional leaves.

3.6.4 Absolute size

Mathematically, shape should be independent of size, but in living systems size matters. On one hand, size is always limited by physical laws but on the other hand it can still cover a few orders of magnitudes. When studying shape change in real systems, changes in size also need to be monitored.

In this dataset it is difficult to identify a trend with respect to the area: in most leaves, cell size increases in the first few centimetres above the ligule — cells close to the ligule are still expanding. It varies where the cell area becomes maximal: in some cases, cells were already quite large in the first segment. Cells towards the tip of the leaf — the oldest cells in the given leaf — are usually smaller than the ones towards the middle. Cell size in the young leaves (7–9) of the 4-week-old seedling are slightly smaller on average.

3.6.5 Shape change in time

As described in Section 3.4, cell shapes change as the plant grows older and also along the leaf, with a slight variation from leaf to leaf. Figures 3.8 and 3.9 show the position of the cells in the two-dimensional morphospace

defined by aspect ratio and solidity. Different panels correspond to seedlings aged 1, 1.5, 2, 2.5, 3 and 4 weeks. Data points are coloured based on their absolute distance from the ligule. The cell shapes inserted in the last panel of Figure 3.9 illustrate (very approximately) the morphospace: along the x axis, cell shape changes from elongated to round and along the y axis, cell shape changes from wiggly to smooth.

The migration of the point cloud from panel to panel confirms that morphogenesis does indeed happen: the shape characteristics change. By observing the values, we can deduce more — decreasing values of aspect ratio (shifting left) indicates that the shape change is due to elongation and decreasing values of solidity (shifting down) signifies deepening undulations.

At week 1, cells occupy a small region of the morphospace and are very confined along the vertical axis — outlines are only slightly or not at all undulating at this point. Aspect ratio varies from round cells to approximately 1:3. It is worth noting that the average aspect ratio decreases as the distance from the ligule increases: mature cells are more elongated. As shown by the next three panels, during the next ten days of the plant's life, the cells keep elongating up to an aspect ratio of 1:8 — the longest cells are more than twice as long as the longest ones at week 1. Simultaneously, the depth of undulations increases slightly, to a solidity of 0.75–0.80 — about 1/8th of the width of the cell. Regarding location, it appears that elongation is still on average higher as we go further from the ligule.

Looking at the panels for 1.5, 2 and 2.5 weeks, it appears that on average there is a monotonous relationship between solidity and the aspect ratio until we reach the third week. Using Spearman's test, we obtain the following values for ρ^2: 0.1467 (1 week), 0.4225 (1.5 weeks), 0.5903 (2 weeks), 0.4667 (2.5 weeks) and 0.3648 (3 weeks). The correlation is strongest for the 2-week-old plant, and near-significant for 1.5 and 2.5 weeks. In fact, the relationship for these three samples looks approximately linear. Linear fitting yields the following equation for the 2-week-old data set: $y = 0.787 + 0.204x$, with Pearson's r^2 being 0.5370.

An interesting outlier appears around the third week in the seedling's life: data points with very low solidity are present in the panels for 3– and 4-week-old plants. The presence of such "stray" points means that at this location the rate of lobe growth increased compared to the elongation rate and the lobe depth approximately doubles. This happens in the regions

furthest from the ligule as indicated by the pink colour.

This transition at 3 and 4 weeks mark the transition between juvenile and adult phases. It has previously been observed that mature cells in the adult phase develop deeper undulations [112]. From the data we can conclude that mature adult-type undulations appear on leaf 7 of a 3-week-old plant at 22 cm from the ligule. Leaves 7 and 8 of the 4-week-old plant already show mature adult cell shapes around 8 cm from the ligule.

According to the relationship between solidity and undulation depth derived above, the mature adult cell shape is characterised by an undulation depth of $\approx$ 30–40% of the total width. In contrast, mature juvenile cell shapes typically display an undulation depth of 20% or less.

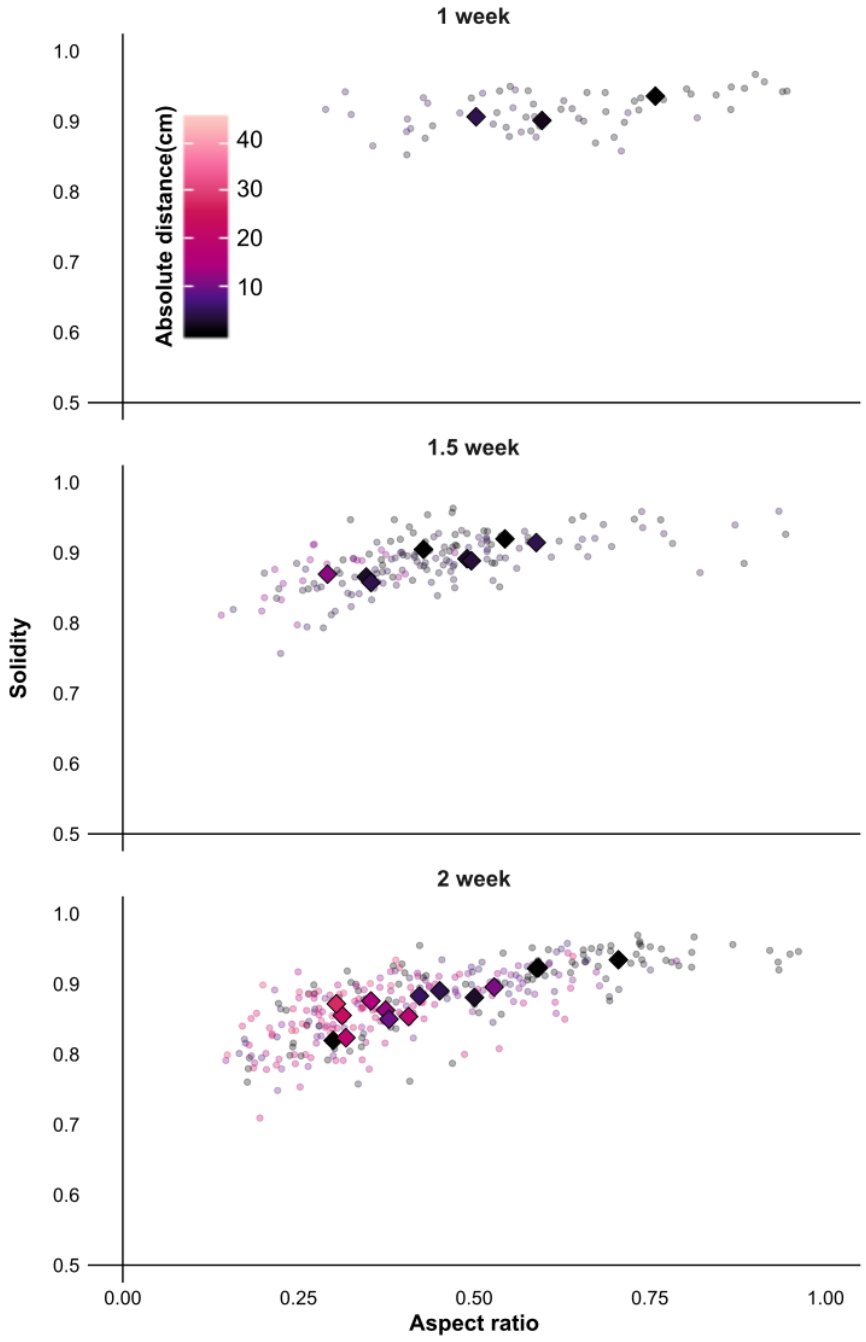


Figure 3.8 – Morphogenesis of maize cells
Solidity vs aspect ratio for seedlings aged 1, 1.5 and 2 weeks. Figure 3.9 shows solidity vs aspect ratio for seedlings aged 2.5, 3 and 4 weeks. The colour of data points corresponds to the location on the leaf (absolute distance in cm). Circles mark individual cells, diamonds shows the sample average. Example cells are included as illustration in the last panel of Figure 3.9 — dark grey: aspect ratio = 0.987, solidity = 0.968; light grey: aspect ratio = 0.113, solidity = 0.906; violet: aspect ratio = 0.570, solidity = 0.605; pink: aspect ratio = 0.186, solidity = 0.595.

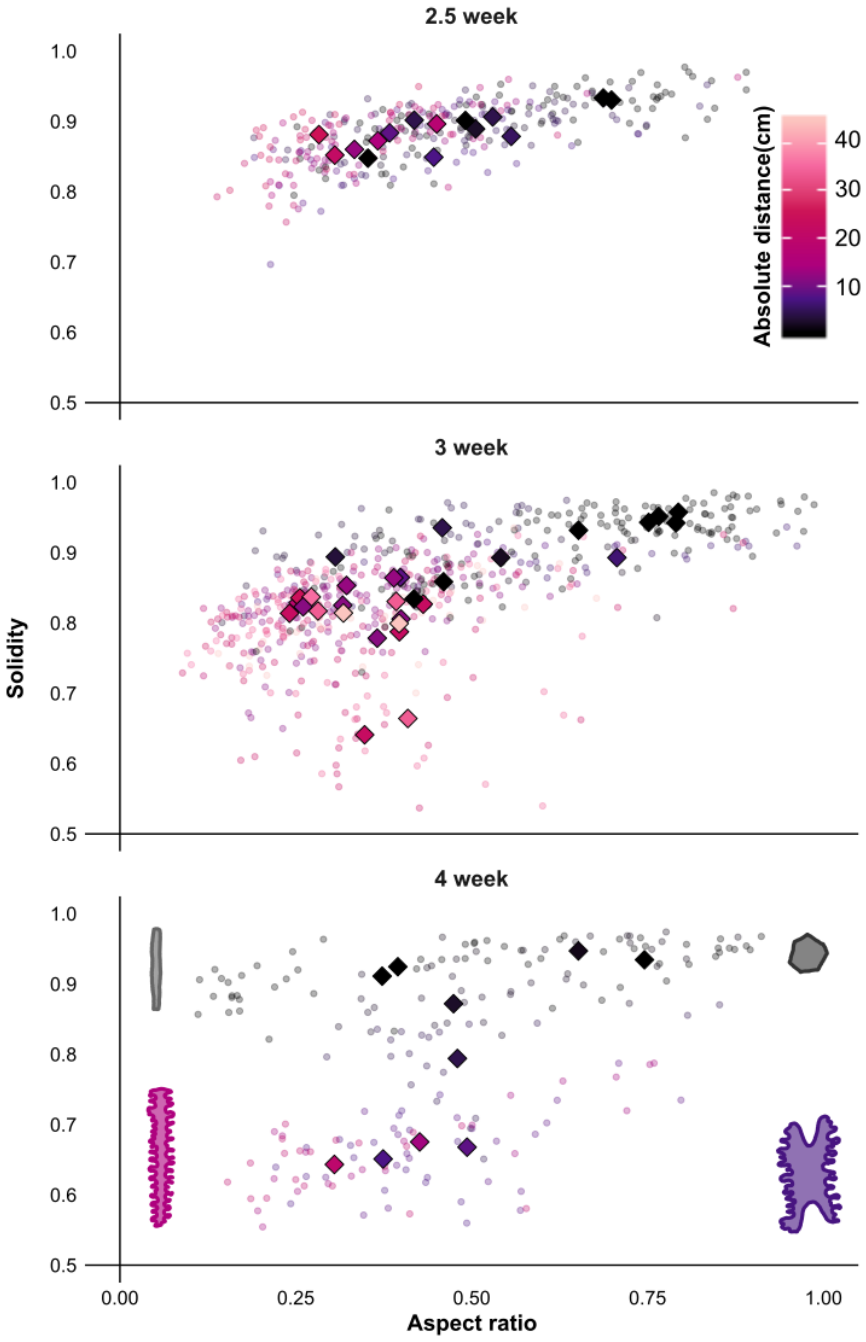


Figure 3.9 – Morphogenesis of maize cells
Solidity vs aspect ratio for seedlings aged 2.5, 3 and 4 weeks. Figure 3.8 shows solidity vs aspect ratio for seedlings aged 1, 1.5 and 2 weeks. The colour of data points corresponds to the location on the leaf (absolute distance in cm). Circles mark individual cells, diamonds shows the sample average. Example cells are included as illustration in the last panel — dark grey: aspect ratio = 0.987, solidity = 0.968; light grey: aspect ratio = 0.113, solidity = 0.906; violet: aspect ratio = 0.570, solidity = 0.605; pink: aspect ratio = 0.186, solidity = 0.595.

3.7 Fourier analysis

As the undulations of the mature cells show some periodicity, it felt justified to try elliptic Fourier descriptors. This was also implemented in the Momocs package. The elliptic method was used and the first 40 harmonics were determined. Figure 3.10 shows the reconstruction of a sample cell from the mature adult set — even with the first 30 harmonics, some of the undulations disappear from the reconstruction.

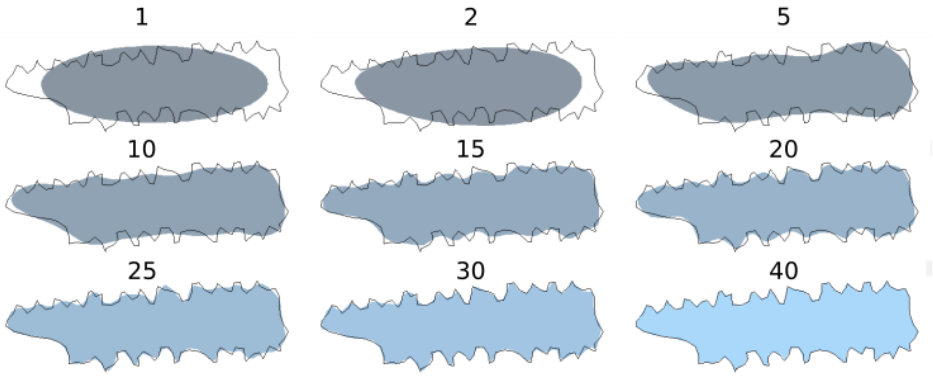


Figure 3.10 – Cell shape reconstruction from the first n Fourier coefficients
$n = 1, 2, 5, 10, 15, 20, 25, 30, 40$

PCA was carried out on the Fourier coefficients. It was found that the first principal component (PC1) covered 81.4% of the variability in the data, with PC2 and PC3 accounting for 2.86% and 2.07% respectively. Contributions of the first four principal components to the shape are shown on Figure 3.11. PC1 accounts for elongation and the convexity/concavity of the long wall. PC2 is responsible for rectangular distortions. PC3 introduces an asymmetric feature at one end, PC4 corresponds to a shear-like deformation. Outline undulations are not present in the first four principal components.

Data from maize cells is shown on Figure 3.12 in the morphospace defined by the principal components. Data was grouped into five classes based on absolute distance along the leaf: closer than 2 cm (red), 2–5 cm (purple), 5–15 cm (blue), 15–30 cm (teal), further than 30 cm (green).

Clusters are heavily overlapping, most of the separation is along the PC1 direction. This corresponds to the already demonstrated variation of

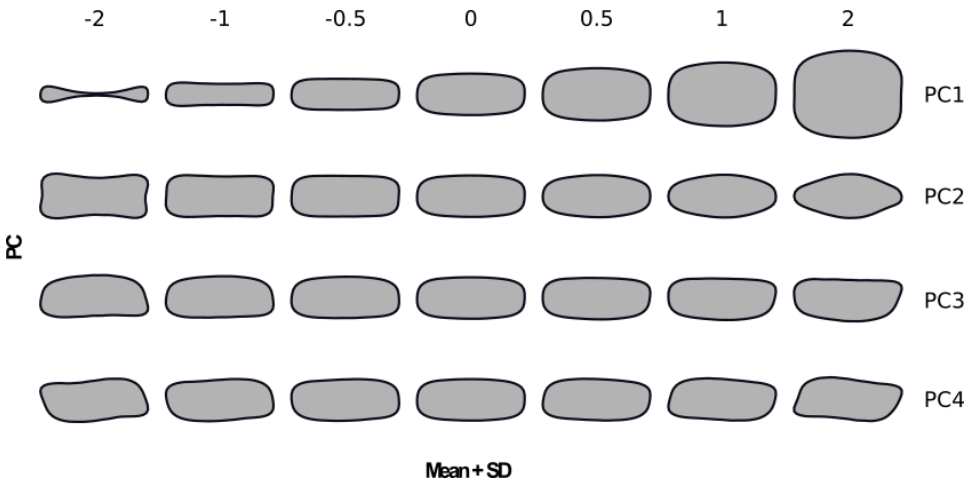


Figure 3.11 – Contributions of the first four principal components to the shape— shapes illustrating the mean shape and shapes ± 0.5, 1 and 2 standard deviations away from the mean shape along the first four principal components.

elongation along the leaf: Spearman's ρ for correlation between the absolute distance and PC1 is -0.5137 and 0.9915 for correlation between the aspect ratio and PC1. The contribution of PC2 to the separation of points can be interpreted as elongated cells having non-convex (undulating) edges and cells with an aspect ratio close to 1 have convex boundaries (round/squareish). PC3 contributes little to the separation of classes.

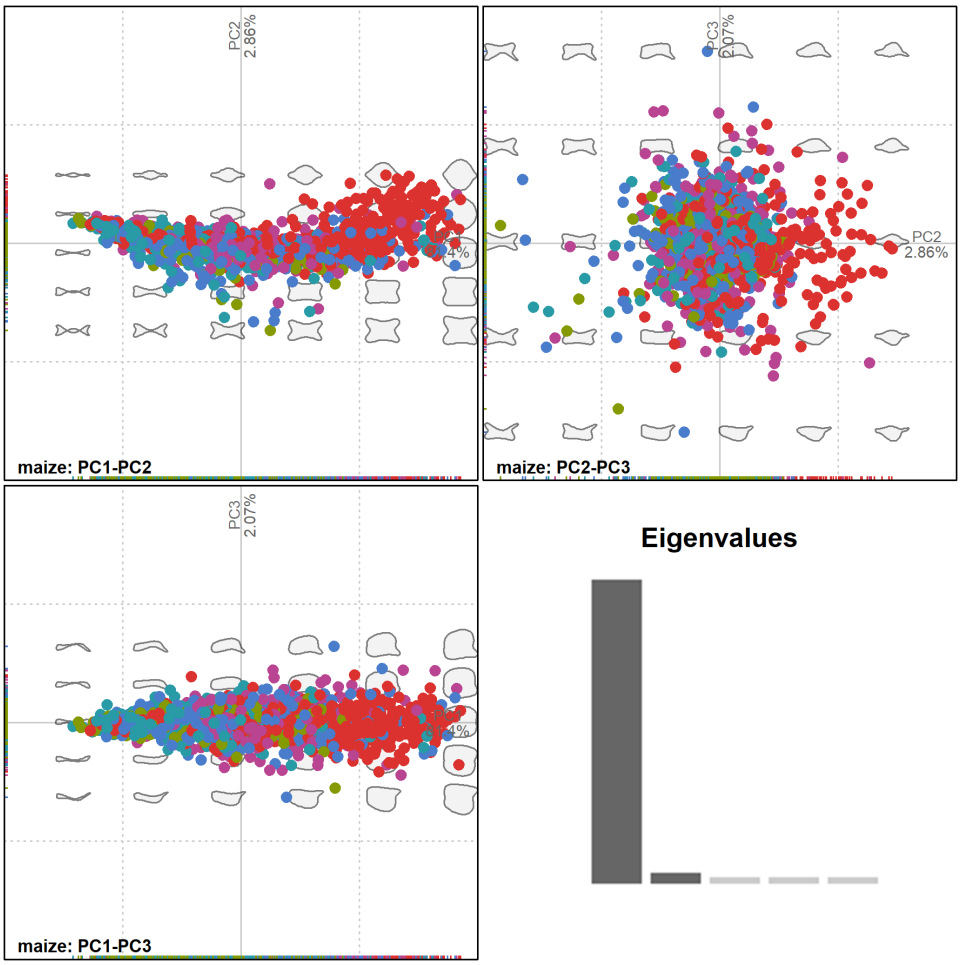


Figure 3.12 – **Maize data as captured by Fourier PCA**
top left PC1–PC2 morphoplane, **bottom left** PC1–PC3 morphoplane, **top right** PC2–PC3 morphoplane, **bottom right** Eigenvalues of the first 5 principal components are 81.4%, 2.9%, 2.1%, 2.0% and 1.7%. Data grouped based on absolute distance along the leaf: closer than 2 cm (red), 2–5 cm (purple), 5–15 cm (blue), 15–30 cm (teal), further than 30 cm (green)

3.8 Discussion

Sampling One weakness of the dataset is the distribution of sampling locations. Initially, five 1-cm-wide segments were chosen for imaging at equal distances along the leaf but this does not give a good coverage of the zone where shape formation happens. The reason for this sampling was due to practical issues: the time taken to acquire an SEM image is limited by the time necessary to pump down the chamber to vacuum pressure after inserting the sample and raising the pressure again before changing samples. Another limitation on the amount of samples was manual segmentation: the complex shape of maize cells means that it can take up to an hour to segment one image. I originally chose to sample at equally placed locations on the whole length of the leaf as it can be interesting to see the variation between mature and juvenile cell shapes in the leaf and when it was obvious that sampling sufficiently densely along the whole length will not be feasible, I chose to increase the distance between sampling location, which, in retrospect, was suboptimal.

Imaging Imaging the cell shapes at a resolution high enough to capture the details of the lobes while fitting entire cells into an image also poses a challenge as the mature adult cell shapes appear. In order to improve image quality and potentially obtain images that can be segmented automatically, confocal microscopy was also attempted.

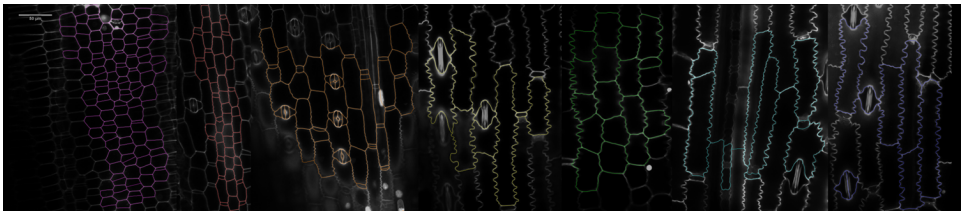


Figure 3.13 – Confocal images of maize cells as segmented automatically — Images taken from consecutive 1 cm segments of leaf 7 of a 2.5-week-old seedling (first 7 cm from the ligule). Different colours correspond to different locations. Scalebar marks 50 µm

Confocal images are significantly easier to use as they have better contrast: for most images, simple thresholding is often enough to get the outlines. Figure 3.13 shows automatically segmented confocal images of maize cells in different stages of development. However, obtaining sharp confocal

images of maize remains challenging: the waxy surface of the maize leaf prevents liquid penetrating the tissue. Propydium iodide was used to stain the cell walls and it was found that using too high concentrations or soaking for too long can lead to staining the nuclei as well. Eliminating this from the images requires a more thorough and less automatic pre-processing.

The third dimension A further problem arises when imaging a well-stained maize leaf under a confocal microscope: the lobes are only present towards the top surface of the epidermal cells. At the bottom of the epidermal cell layer, the cell cross-section is entirely lobeless. As one moves the focus along the z-axis from the bottom of the epidermal cell layer, the lobes first appear and become more pronounced towards the surface. It would be tempting to take images focused onto the upper surface, however the propidium iodide also stains the upper cell wall of the epidermal cells, reducing the contrast around the lobes.

Moreover, because the lobes are only present in an upper part of the cells, the sample has to be very flat or meaningful cell outlines can only be extracted by taking a stack of images of each area selected, where each image is focused at a slightly different z-stack; reconstructing the 3D shape of the epidermal cell layer from this data and finding the surface of the leaf in order to acquire a correct 2D shape. This problem does not appear when using SEM images.

Shape descriptors In this experiment, it was found that geometric descriptors outperformed Fourier descriptors. Even though care was taken to include a sufficient number of terms ($n = 40$) for the Fourier representation, outline undulations were suppressed by PCA. The reason for this is that the contribution of undulations is spread out amongst the higher harmonics instead of being captured by just one. Because the relative contribution of higher harmonics decreases as the number of harmonics increases, Fourier descriptors intrinsically focus on the base shape of the cells rather than the outline. In this case, the weak clustering observed in the PC1-PC2 plane is due to the variation in elongation. Meanwhile, a better representation was provided by employing just two geometric variables: aspect ratio and solidity. This method was successful exactly because solidity captures the presence of undulations globally, indifferent of the structure. In the simple

geometry of the maize cell, variations in solidity are almost always due to the presence of undulations in the cell wall.

3.9 Conclusion

In this experiment, a large amount of data representing different stages of the developing maize cell was collected — although the distribution of the samples could be more optimal. Even though cell tracking through different phases of development was not possible, the structure of the maize leaf allowed us to follow the morphogenesis. Initially, lobe depth and elongation are linked: cells further away from the ligule elongate and slowly grow deeper lobes. Qualitative shape analysis agreed with the previous findings, namely that maize cells go through an elongation phase followed by a maturation phase and that the mature shape is indeed different in juvenile and adult leaves. A first quantitative analysis confirmed that geometric morphometric descriptors coupled with PCA are capable of capturing the variation in the dataset sufficiently to separate cells along two quasi-orthogonal directions that roughly correspond to the qualitative descriptors. This was followed by the comparison of an even more streamlined geometric and a standard elliptic Fourier analysis, of which the first one was not only much simpler but also more efficient in distinguishing cell shapes.

Results of these calculations confirmed that shape maturation happens in the first 5–10 cm of the leaf and quantified the difference between the mature cell shape in adult and juvenile leaves: mature juvenile cells stay at a solidity value greater than 0.8 while mature adult cells go as low as 0.65. This change corresponds to lobes nearly twice as deep (in proportion to the full width of the cell). Plotting the data in the solidity (lobe depth) versus aspect ratio (elongation) morphospace, we can follow the (statistical average) morphogenesis of the cells and observe that elongation is accompanied by a proportional drop in solidity. An interesting finding is that the rate of elongation and the rate of lobe growth appear to be linked until the seedlings reach 3 weeks. The rate changes when the first transition phase leaves appear with mature adult cells.

Chapter 4

Variety in vascular plants

4.1 Motivation

After identifying how much maize cells can differ at different stages of the plant's development and in possession of tried and tested numerical tools, the curiosity grew: how different are the pavement cells of other species?

Looking at model organisms at hand, *Arabidopsis thaliana*, tobacco and tomato all develop jigsaw-like cell shapes, although the complexity of lobes differ [19, 122]. Maize also shows deep undulations in the anticlinal wall of pavement cells. *Brachypodium dystachion* is an outlier: its cells are elongated, like those of maize, but their outline is perfectly smooth. One begins to wonder: is this observation representative? Are undulating cell shapes really so common in the leaf epidermis of vascular plants?

Literature in this respect is scarce: the available analyses for species other than *A. thaliana* are almost all qualitative [10, 13, 123, 124, 125]. Quantitative cell shape analysis are routinely done in plant fossils [91]. To the best of my knowledge, there are no digital databases available that collect epidermal cell shape data or images.

Between older publications hypothesising about the origin of undulations employing physical consideration and modern studies investigating the complex biochemical regulation of lobe outgrowth primarily in *Arabidopsis* and to a lesser extent, maize, the available information about cell shapes in other species is little and unstructured. Eudicots and ferns are said to have un-

dulating cell walls [10, 126, 127], while monocots — with the exception of grasses — have straight outlines. The cells on the abaxial side are claimed to show more undulation than on the adaxial side [13].

These observations came from a sporadic sampling and to confirm their validity, further studies are necessary. Thus the idea for a project bridging this gap by surveying a large number of species and quantifying the observed cell shapes was born. With such a collection at hand, we could also examine how the cell shape varies with respect to other features, and how it correlates with phylogenetic distance.

4.2 Aims

In collaboration with Madelaine Bartlett and her lab (UMass Amherst), we first set out to establish a database of cell shapes from a large number of species and quantitatively describe these using either Fourier or geometric descriptors and principal component analysis, so the results can be interpreted in a phylogenetic context.

As the leaf is shaped by the growth of its cells, it seemed promising to also record leaf shape to study the correlations between cell shape and leaf shape. Cell shape characteristics from the ab- and adaxial sides were compared whenever data from both was available, with the intention to examine whether it is true in general that adaxial cells are less undulating.

Because this project was the result of a collaboration, contributions are detailed in Section 4.3. Experimental methods are presented in Section 4.4. Section 4.5 presents the quantitative analysis of cell shapes, including a comparison of the performance of the Fourier and the geometric descriptors and detailed discussion of the results by clades. Section 4.6 studies the underlying phylogenetic correlations. Section 4.7 compare the cell shape with the leaf shape, Section 4.8 analyses the differences between ad- and abaxial sides.

4.3 Contributions

Species from Amherst were collected by Grace Pisano, Rebecca Goldberg and Jeffrey Heithmar under the supervision of Madelaine Bartlett. Imaging and segmentation of these samples was done by Grace Pisano, who then

shared the single cell images with me for analysis. Species from Cambridge were collected, imaged and segmented by myself. The analysis presented in Sections 4.5, 4.7 and 4.8 is entirely my own work. Section 4.6 discussing phylogenetic correlations is based on the work of Madelaine Bartlett, who carried out the calculations on the full dataset and created the visualisation. This part is included in this thesis because it was an integral part of this collaborative project. Results from Sections 4.5–4.8 were written up for publication and the manuscript is currently under review [128]. These parts of this chapter are adapted from this paper and may therefore show full or partial verbatim agreement with [128].

4.4 Experimental methods

4.4.1 Plant material

Fully expanded adult leaves were collected from healthy plants grown in one of two locations between September 2015 and December 2016: The Botanic Garden of the University of Cambridge (132 species) or the Natural History Collection of UMass Amherst (98 species). For a list of species including location, see Appendix C. One leaf was used per species and cell images were taken from one segment of the same leaf whenever possible.

4.4.2 Sample preparation

Two methods of sample preparation were used; First, when possible, epidermal peels were removed from the adaxial side of the leaf. When this was not possible, the abaxial side was attempted. Secondly, when peels were unachievable, I followed a dissection and maceration protocol obtained from Siobhan Braybrook.

Roughly 5-by-5 mm asymmetric trapezoids were cut from the leaves, near the midrib, halfway along the length. The asymmetric shape allows keeping track of adaxial and abaxial sides through the several-day-long process. These pieces were placed in multi-well plates and soaked in approximately 1 ml of a 1:7 mixture of acetic acid and 100% ethanol overnight at 4 °C, stirred at 50 rpm. The following day the solution was removed and samples were washed three times for 10 minutes. After the last wash, water was replaced by 1 ml of 1M NaOH solution and left to stand for 24 h at room

temperature, without stirring.

Following this, the samples were washed again as before, and the solution was replaced by 1 ml of a solution containing 250 g chloral hydrate dissolved in 100 ml of a 1:2 mixture of glycerol and water. The samples remained in this solution for 3–5 days, until they became fully transparent. When the clearing finished, the samples were washed again as before and stored in water.

4.4.3 Staining and imaging

Samples were stained with 0.1% toluidine blue in water overnight, mounted on glass slides and covered with a cover slip. Images were acquired at 200×, 400×, 700× or 1000× magnification using a Keyence VHX-5000 digital microscope (Cambridge; Keyence UK & IL) or an Axioplan microscope (Amherst; Zeiss, DE). The lowest magnification where every detail of the outline was clearly visible was used. Whenever possible, images were taken from both sides of the sample, at the same magnification. Images were taken at the default screen resolution 1600×1200, corresponding to resolutions of 0.97, 0.51, 0.30 and 0.21 µm depending on the magnification and saved in .tif format to avoid errors resulting from file compression.

4.4.4 Segmentation

Segmentation was done manually, using a freely available image editor (GIMP [116]), resulting in a black-and-white image of cells. The coordinates of the cell outlines were extracted in MATLAB [118], using functions included in the basic distribution. For every species, 30 cells were included in the final dataset, from both the adaxial and abaxial sides when available. The final dataset counted 10260 cells.

Following the difficulties with the segmentation in the maize project (see Section 3.3.3), the experimental procedure in this part was designed with the intention to capture images suitable for automatic segmentation. SEM imaging was ruled out because imaging at sufficiently high resolution compromised the stable running of the software and the problems with the propidium iodide staining in the case of maize made confocal microscopy less appealing as we anticipated that plants with similar leaves may also stain badly.

Toluidine blue staining of the epidermis in most cases gave very good contrast but often the staining was uneven: much stronger closer to the edges. Trichomes often stained strongly and blocked the outline of several cells. Dust grains sticking to the surface were also visible sometimes. These artefacts made automatic segmentation based on watershed algorithms impossible.

Using an adaptive segmentation software, ilastik [129], moderate success was achieved: after a lengthy training period, the machine learning algorithm could eventually provide an acceptable result, although the pixels on the outline were often misclassified. Unfortunately, while toluidine blue was found to be a good stain for most species, the quality, the intensity and the contrast of the colour varied quite significantly. This meant that the training procedure had to be repeated for every species separately. At this point, manual segmentation became quicker and more accurate.

4.4.5 Leaf shape

To be able to analyse the leaf shape, leaves were flattened and scanned in front of a white background at a resolution of 300 dpi. These images were first binarised using a simple threshold and the outlines were then extracted using MATLAB [118].

4.4.6 Shape analysis

Cell outlines were used to calculate traditional morphometric descriptors: absolute area in μm^2 for cells and mm^2 for leaves, aspect ratio, circularity and solidity as defined in Chapter 2 and to extract elliptic Fourier descriptors. Descriptors were calculated and PCA was carried out using the Momocs package in R [90]. For PCA on the set of geometric descriptors, area and circularity were log-scaled in order to better compare them to aspect ratio and solidity, which are bounded quantities ranging from 0 to 1.

Data points were divided into 5 categories depending on where the species belongs, these categories were: ferns (4 species), gymnosperms (19 species), early diverging angiosperms (consisting of one species from the Austrobaileyales, one species from the Chloranthales and five Magnoliids), monocots (64 species) and eudicots (137 species).

Correlations between traditional metrics were examined in R using the

stats package (included in the standard distribution) [119]. Correlations were tested using Spearman's ρ to test for a monoton relationship and least square fitting was also used to test the linearity of the relationship. Comparison of distributions were also performed in R using Mann-Whitney and Kruskal-Wallis tests between different groups.

4.5 Cell shapes

Even through the collection phase of this project it was apparent that this dataset is a lot more varied than the maize one. Elongation and undulation depth here appeared to vary independently and in a more complex manner than with maize cells. Nevertheless, outline quality and elongation seemed to remain important factors. Geometric morphometric descriptors have previously been proved to capture these two characteristics reasonably well (see Chapter 3). Elliptical Fourier descriptors were also tested as this is currently the most popular method for shape quantification. Both of these methods were followed by a further reduction of variables through principal component analysis.

4.5.1 Fourier descriptors

When using elliptic Fourier descriptors, the standard way to calculate a cutoff point is to determine the number of harmonics required to reconstruct a shape that captures a given proportion of the area of the shape. In Momocs's `efourier` function the default value is 99% although it can be adjusted.

This area-based method, however, does not prove appropriate for capturing the undulation of the cell walls. While only 11 harmonics were found to be sufficient to capture the area with 99% precision, visual inspection shows that this remains quite inaccurate in terms of the outline. Figure 4.1 shows the reconstructions of two cells: the first has a fairly complex shape with a solidity of 0.6 (in the bottom 3% of the scale) and the other is the most complex shape with the lowest solidity, 0.4.

The cutoff point was therefore determined in a different way. Figure 4.1 shows shape reconstructions of a cell with a solidity of 0.6 and the cell with the lowest solidity in the entire dataset. Only 2.3% of the cells have a solidity that falls within the range defined by these two cells. By visual inspection, using 20 harmonics gives a good representation of the first cell

and an acceptable one of the second. Because this dataset consists of more than 10000 cells, every added harmonic meant a significant increase in computation time. Therefore, the cut-off point was set to 20 in this analysis.

Using a comparison between the perimeter of the reconstructed shapes and the perimeter of the original shapes to determine the cut-off point (similar to the original area-based method but more appropriate for studying outlines) was also contemplated. However, implementing this would have taken significantly longer than the visual determination described above. For this reason, the visual determination method was selected.

After calculating the elliptic Fourier descriptors, the principal components were identified. Figure 4.2 show the contribution of the first four principal components to the shape. The first principal component represents 55.17% of the variability in the data and is again dominated by the elongation of the shape, with a mild influence of symmetric convex or concave walls. PC2 and PC3 (11.11% and 7.89%) capture symmetries in the dataset: rotational symmetry (PC2) and inversion (PC3). PC4 (3.85%) represents a convex to concave change on one of the long sides. Again, the outline undulations seem to be lost and only the elongation of the base shape is captured reasonably well.

To further test the universality of outline suppression with the standard elliptic Fourier method, two monocot species were selected from the sample set: *Danthonia californica*, whose cells are elongated, with a smooth, lobeless outline; and *Spathiphyllum wallisii*, whose cells remain close to 1 in aspect ratio but their outline is undulating. Cell shape data from these two species were handled separately and the same method of calculating elliptic Fourier descriptors and applying PCA was performed. Results of this test are shown on Figure 4.3. Cell shapes in pink (I.) are some of the original cell shapes as segmented, the bar plot shows the eigenvalues of the principal components (II.) and the plot shows the location of the data points associated with individual cells on the PC1–PC2 morphospace in each case. The shapes associated with different regions of the morphospace are depicted in grey.

In the case of *Danthonia*, the original cells (pink) look very similar to the reconstructions in the PC1–PC2 morphospace. This means that for cells like these, the Fourier method gives a good description. In fact, given the very high weight of the first principal component (88.79%) and the small spread of points along the PC2 axis, just the first component would be enough to

distinguish between cells in this sample, presumably based on elongation as the outline is smooth.

The situation with *Spathiphyllum* is different: here the reconstructed (grey) cells, while similar to the base shape of the original cells (pink), are clearly missing the characteristic undulations of the outline. The weight of the principal components is also much more evenly distributed. In fact, the first two principal components cover less than 50% of the variability. For complex cell shapes with undulating walls, Fourier decomposition in its standard form is not a good representation as the outline features — the particular property we set out to study — are lost. In contrast, if one wanted to compare the base shape of cells but ignore the variation in the outline, the Fourier method would be a great choice.

Returning to the full dataset with these limitations of the technique in mind, we can understand why the separation of clusters on Figure 4.4 is poor. Eigenvalues of the first five principal components and the contributions of the first ten harmonics to PC1 and PC2 are shown in the top left corner. PC1 is nearly entirely composed of the first harmonic (the elongation) and PC2 has minimal contributions beyond the fourth harmonics — yet again, higher harmonics and therefore undulating outlines are suppressed.

Gymnosperms (purple) and monocots (pink) are confined to the PC1 axis, meaning that these cells are only characterised by elongation and suppressing cell shapes like those observed in maize. Ferns (blue), early diverging angiosperms (turquoise) and eudicots (green) spread out more and occupy a similar area, indicating that the base shape of the cells in these groups is more irregular than in gymnosperms and monocots but any further distinction is difficult.

4.5.2 Geometric descriptors

Area, aspect ratio, circularity and solidity were calculated for every cell and this data was subjected to principal component analysis. The first two PCs represent 50.37% and 24.80% of the variability in the data, which is more than what is captured with the Fourier analysis. Scores in the PC1–PC2 morphospace are shown on Figure 4.5, along with the projection of the geometric quantities onto this plan in the top left corner.

We see that the area (A) and the aspect ratio (AR) vary oppositely. As the area grows, the aspect ratio decreases: in this representation, greater

area is linked to greater elongation. Solidity (S) is close to perpendicular to the aspect ratio–area axis. This means that in this morphospace, solidity can vary independently of the aspect ratio and the area: outline undulations are finally captured. Circularity (C) varies both with the aspect ratio and the solidity: this is a direct consequence of the way it is defined and not necessarily characteristic of the dataset.

Looking at the five groups we see that monocots and gymnosperms spread out more along the axis defined by the aspect ratio and the area, but are not so confined as with the Fourier analysis. Monocots show a bias towards higher solidity values indicating that undulations in the outline are less prevalent, but still present. Gymnosperms cover a similar region but are more evenly distributed. Ferns are located in the lowest solidity quadrant indicating complex cell shapes, although there are only four fern samples in our data set. Early diverging angiosperms and eudicots behave similarly here as well: these groups show greater variation in the cell outline than in the cell elongation.

The values for the five groups overlap around the origin of the PC1–PC2 plots. This signifies that the five groups contain similar cell shapes. These are typically high solidity (low undulation) and high to medium aspect ratio (round or moderately elongated) shapes. Multimodularity may occur in ferns and early diverging angiosperms as a result of insufficient sampling (4 and 7 species respectively). However, the two maxima in monocots clearly signifies that apart from the omnipresent round or only moderately elongated cells with smooth outlines, a significant proportion of monocot species exhibit a heavily elongated average shape.

Because aspect ratio and solidity are practically independent in this representation, we can reduce the complexity of the graphs by examining the location of the data points in the solidity–aspect ratio plane. The dataset includes 230 species in 98 families and 49 orders. For ease of visualisation, data was grouped into orders following the suggestions of the APG IV classification [130] and average solidity and aspect ratio values were computed for each order and presented in Figure 4.6. Points are coloured according to the five groups that were established previously.

Example cell shapes are also provided along the plot area. Shapes along the top edge correspond to aspect ratio values of < 0.1, 0.2, 0.4, 0.6, 0.8 and 1, shapes along the right edge correspond to solidity values of 0.6, 0.7,

0.8, 0.9 and 1. The two black lines correspond to an aspect ratio of 0.5 and solidity of 0.75 and divide the data into three groups.

Most of the data is clustered in the top right corner (no to moderate elongation, no to moderate undulations). A close-up of this area is presented in Figure 4.7. Figure 4.7 includes two panels showing the same data but with different groups labelled for better visibility.

Points in the top left corner correspond to highly elongated cells with no to moderate undulations and occupied mostly by gymnosperm and monocot species (Fagales of the Eudicots contains a single species: *Lithocarpus henryi*). The bottom right corner corresponds to highly undulating cells that are not particularly elongated and is populated by fern and eudicot species, save for Laurales comprising a single species, *Laurus nobilis*.

In the following, some results including averages for families or even individual species are highlighted, without illustrations. The full dataset is provided in Appendix C.

Gymnosperms Data was collected from the following gymnosperm families: Cycadaceae: 1 sample, Zamiaceae: 6 samples, Ginkgoaceae: 1 sample, Gnetaceae: 1 sample, Pinaceae: 3 samples, Araucariaceae: 2 species, Podocarpaceae: 1 species, Cupressaceae: 3 species, Taxaceae: 1 species). These cells typically elongate significantly and in some cases (Zamiaceae), the solidity value also drops, indicating more complex shapes. The overall complexity of the outline remains low.

Early diverging angiosperms Data from one *Illicium* and one *Chloranthus* species along with five Magnoliid species was collected. Average maximum cell elongation is around a 1:2 ratio, Piperaceae and *Illicium* species reaching 1:4. Solidity indicates mild-to-moderate undulations except in *Laurus nobilis* (Lauraceae), where the undulations are much stronger.

Monocots Data for Commelinids consist of Arecales, Poales, Commelinales and Zingiberales. Of these, Poales (Poaceae: 10 species, Cyperaceae: 1 species) exhibit very elongated cells with margin undulations. The Commelinaceae family (4 species) also reaches lower solidity values but typically at medium values of elongation. The one species representing another family

from the Commelinales, Pontederiaceae, is different: it develops less elongated cells with shallow undulations (higher solidity values).

Zingiberales (Zingiberaceae: 3 species, Costaceae: 2 species) and Arecales (Arecaceae: 1 species) stay close to maximum solidity, suggesting a lack of margin undulations. Asparagales (Orchidaceae: 4 species, Iridaceae: 3 species, Asphodelaceae: 4 species, Amaryllidaceae: 5 species, Asparagaceae: 8 species) exhibit highly elongated and smooth cell shapes, without any margin undulations. Lower solidity values at very low aspect ratio are artefacts caused by a slightly curved cell shape. In contrast, Liliales (Alstroemeriaceae: 2 species, Smilacaceae: 1 species, Liliaceae: 1 species) exhibit moderate to strong margin undulations but elongate less than the Asparagales. Finally, Acorales (Acoraceae: 1 species) and Alismatales (Alismataceae: 1 species, Araceae: 8 species) develop moderately undulating cells.

Eudicots Eudicots show a diverse range of cell shapes. Some families with more than two genus present show a very wide variation in solidity and mid-to-high maximum elongation. Euphorbiaceae is particularly interesting as one of the three *Euphorbia* species has cell shapes of a much lower average solidity than the other two. In contrast, Oxalidaceae seems homogeneous: all of these three species have moderately elongated cells (length-to-width ratio up to 1:3).

Myrtales (Onagraceae and Lythraceae) and Malvales (Malvaceae) along with Brassicaceae show deeper undulations in the cell outline (solidity below 0.75). Geraniales (Geraniaceae) and Sapindales (Anacardiaceae, Burseraceae, Rutaceae, Sapindaceae) lack undulations except for *Citrus* × *limon*. All species in this group have less elongated cells (typical width-to-length ratio below 1:2).

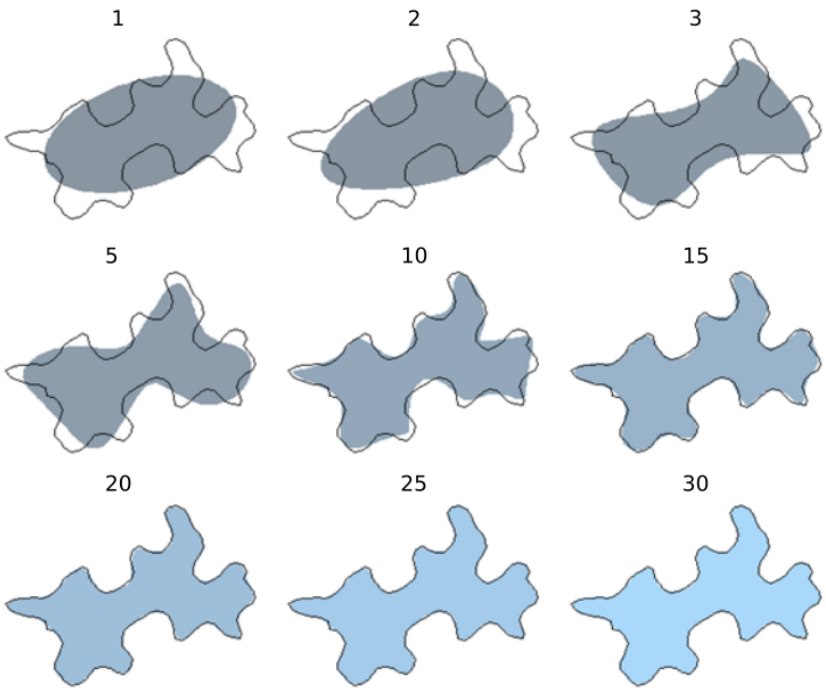
Saxifragales contains 7 species which show a lot of variation in outline character amongst each other. *Paeonia tenuifolia* develops cells with a solidity value close to 0.75. The Crassulaceae family, in contrast, have cells with a smooth outline. Saxifragaceae species show a broad variety of cells in aspect ratio as well as in solidity.

Caryophyllales, with the exception of Cactaceae, develop cells with solidity from as low as 0.5 to 1.0 and not exceeding 1:4 width-to-length ratio in elongation. Cell patterns within families are similar, with Polygonaceae

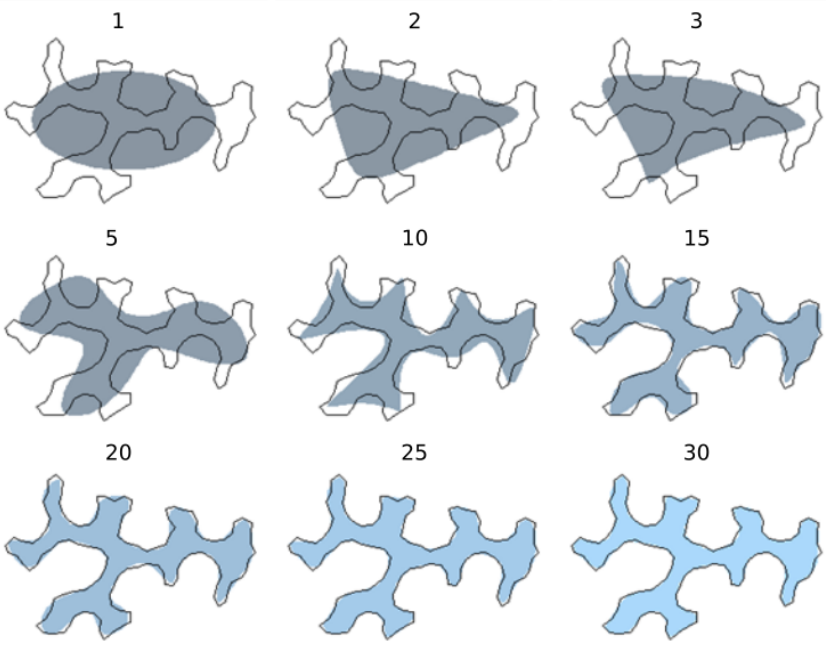
(7 species) showing the deepest undulations followed by Amaranthaceae (5 species) and Plumbaginaceae (2 species) and Phytolaccaceae (1 species) displaying more moderate undulations.

Apiales, Dipsacales and Aquifoliales show deep undulations in the outline on average (lowest solidity around 0.6). The undulation depth of Asterales (Asteraceae and Campanulaceae) cells is slightly lower. As opposed to what was found with the Caryophyllales, where the variation within a sample was quite high, here, although data points for the family as a whole cover a similar range of values, the separation between species in the same family is much stronger. For example in Campanulaceae, *Campanula* species show an average solidity of 0.8 on the abaxial side and 0.9 on the adaxial side, while *Jasione heldreichii* exhibit more undulating cells with solidity just above 0.7 on both sides. Apiaceae species behave similarly: average solidities of the cells of *Bupleurum fruticosum* are 0.87 and 0.93 (abaxial and adaxial, respectively), but *Pastinaca sativa* goes as low as 0.63 and 0.84.

Lamiales show a great variation in cell descriptors, however the solidity appears to be more characteristic of a given species. These species also show a clear separation between ad- and abaxial sides in undulation depth. Solanales (Solanaceae, Montiniaceae) along with the Apocynaceae family from Gentianales also show relatively constant solidity values amongst the cells from the same species, but a wide variation between species. Rubiaceae and Gentianaceae species are more spread out in the solidity direction. Members of the Boraginaceae family are comparatively less diverse in elongation but more so in aspect ratio.



(a) *Fuchsia magellanica*, solidity = 0.6



(b) *Osmunda banksiifolia*, solidity = 0.4

Figure 4.1 – **Fourier reconstructions using the first n harmonics**, $n = 1, 2, 3, 5, 10, 15, 20, 25, 30$

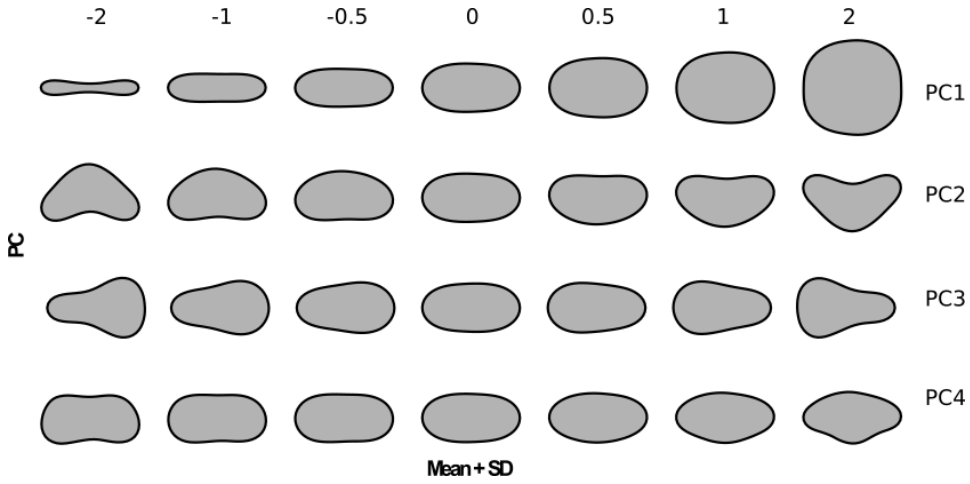
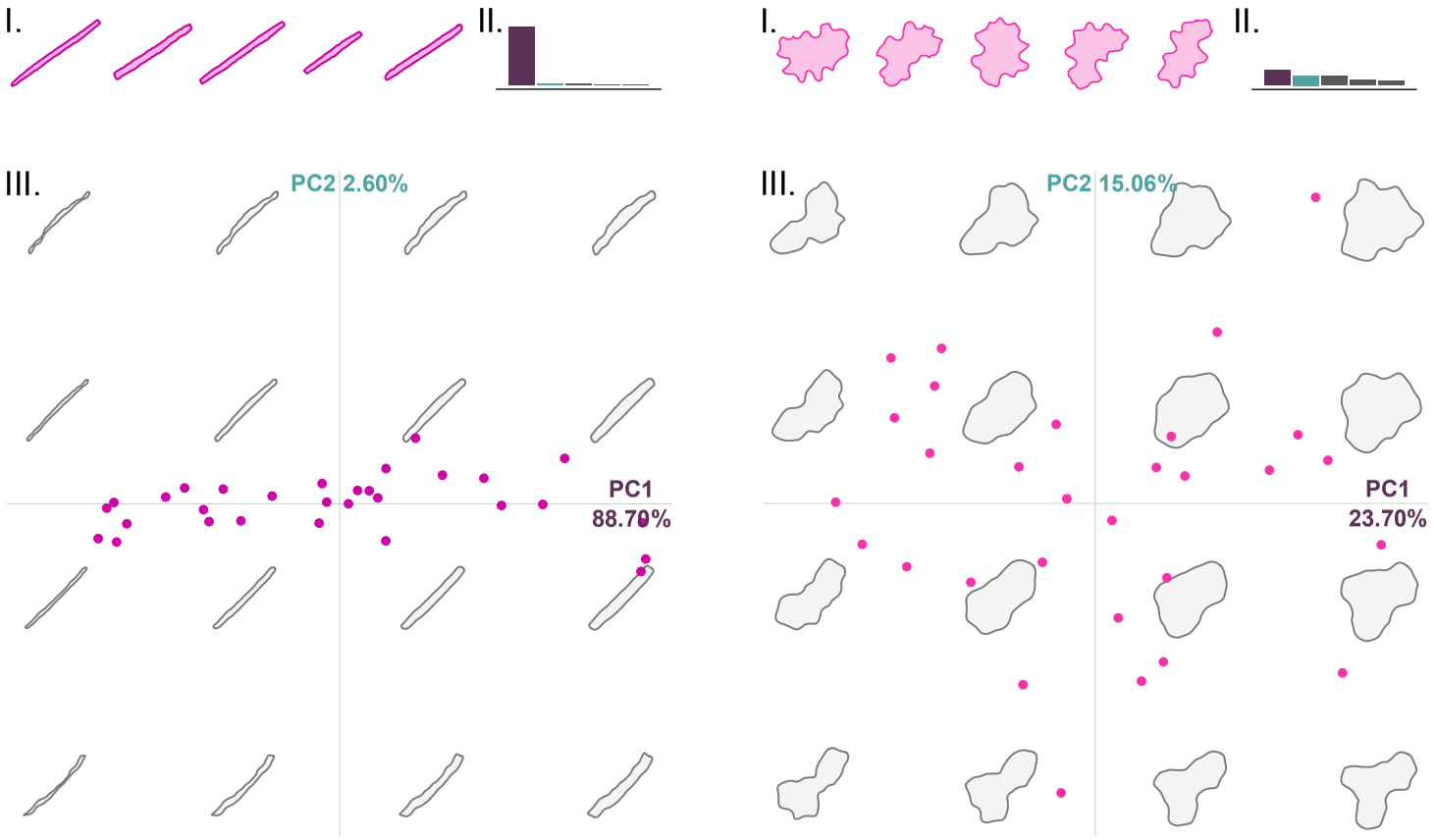


Figure 4.2 – Contributions of the first four principal components to the shape



(a) *Danthonia californica* **(b)** *Spathyphyllum wallisii*

Figure 4.3 – Individual Fourier analyses of two species
I. — some of the original cell shapes, II. — eigenvalues of the first five principal components, III. location of the cells in the PC1–PC2 morphospace

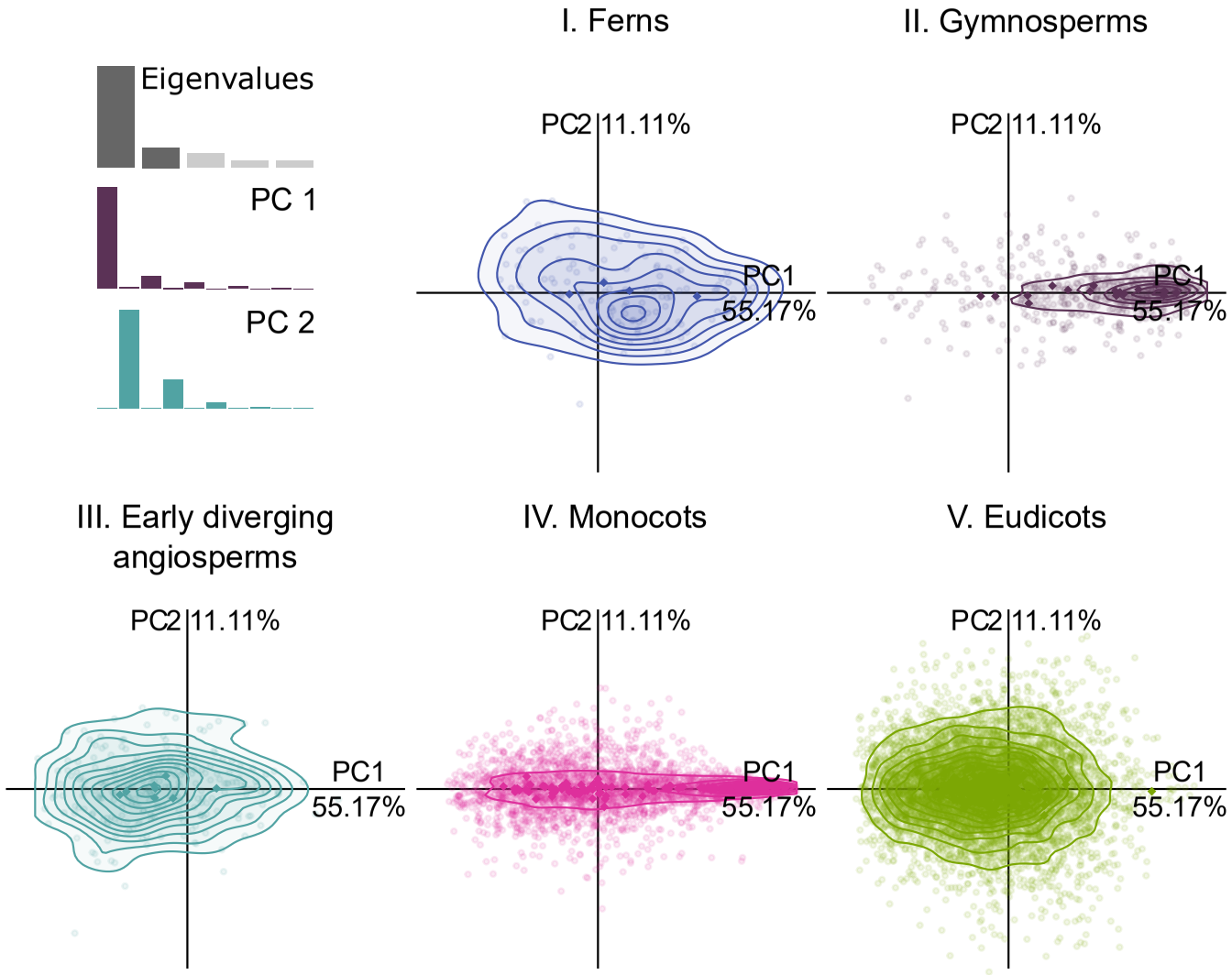


Figure 4.4 – **Fourier shape analysis**
The first panel presents the eigenvalues of the first five PCs and the harmonic composition of PC1 and PC2. The other five panels present the location of Ferns, Gymnosperms, Early diverging angiosperms, Monocots and Eudicots in the PC1-PC2 morphoplane.

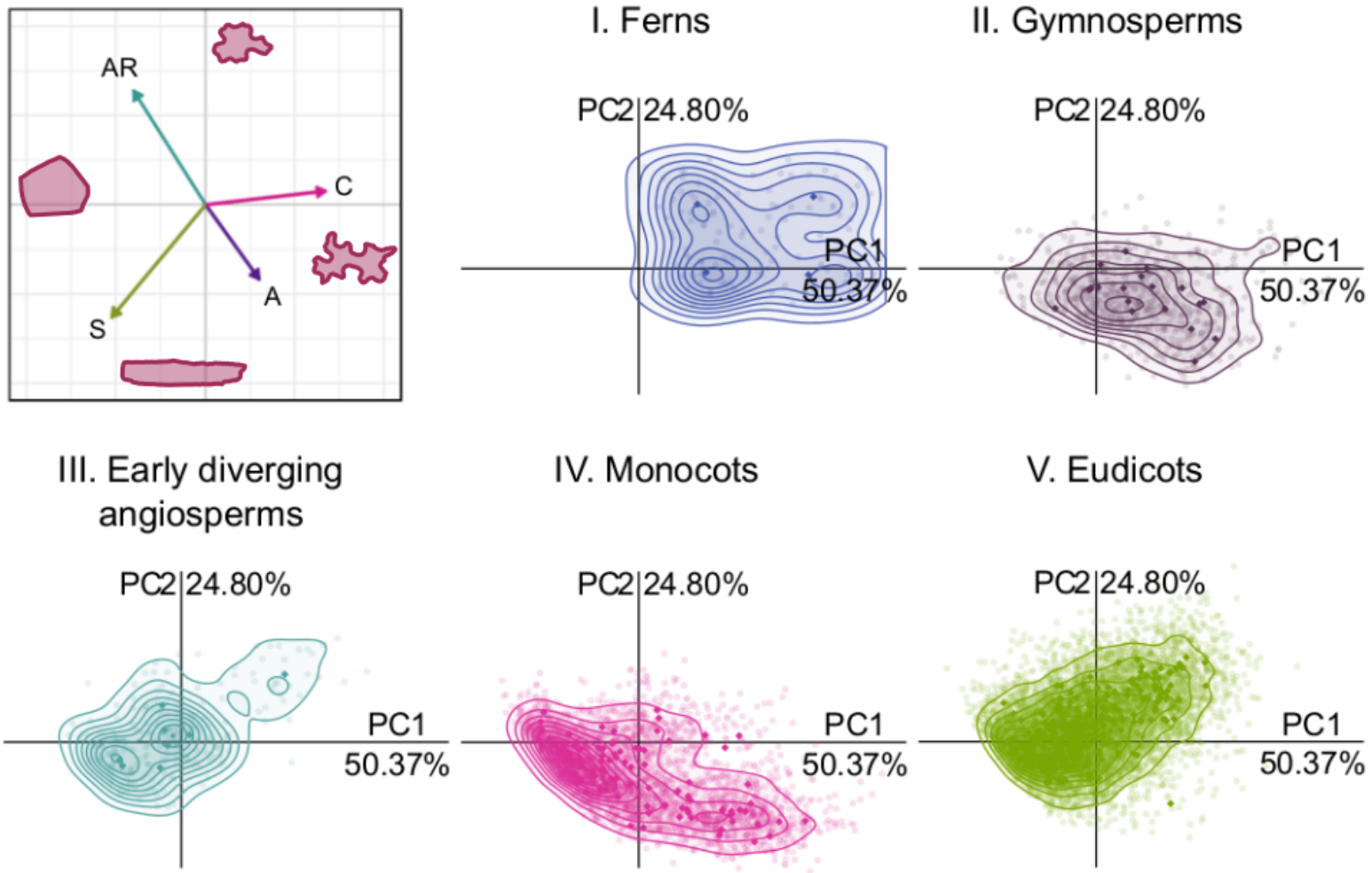


Figure 4.5 – Geometric descriptors
The first panel presents the geometric descriptors in the PC1–PC2 morphoplane: A — area (purple), AR — aspect ratio (turquoise), C — circularity (pink), S — solidity (green). Example cell shapes are provided as illustration. The other five panels present the location of Ferns, Gymnosperms, Early diverging angiosperms, Monocots and Eudicots in the PC1-PC2 morphoplane.

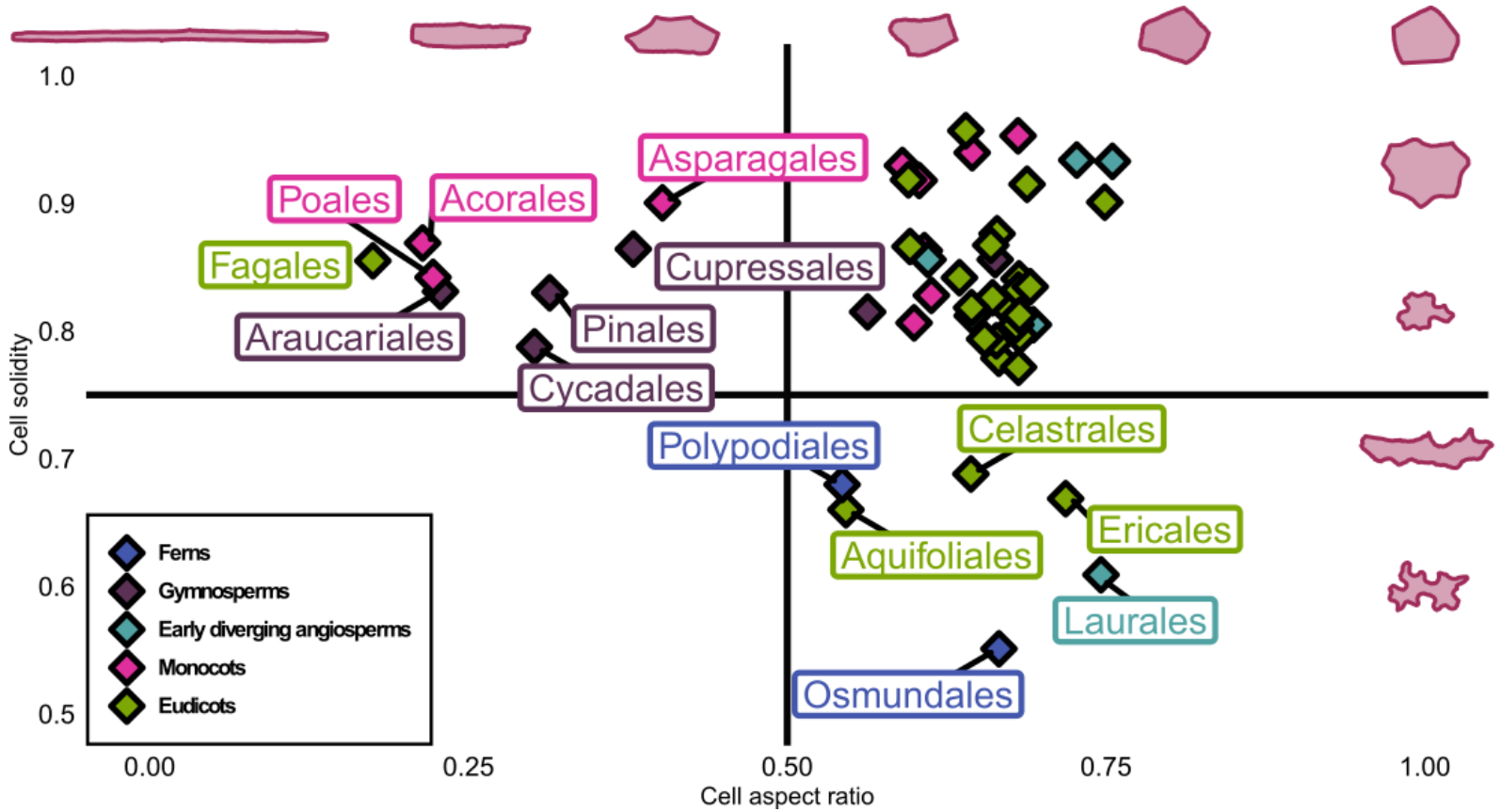
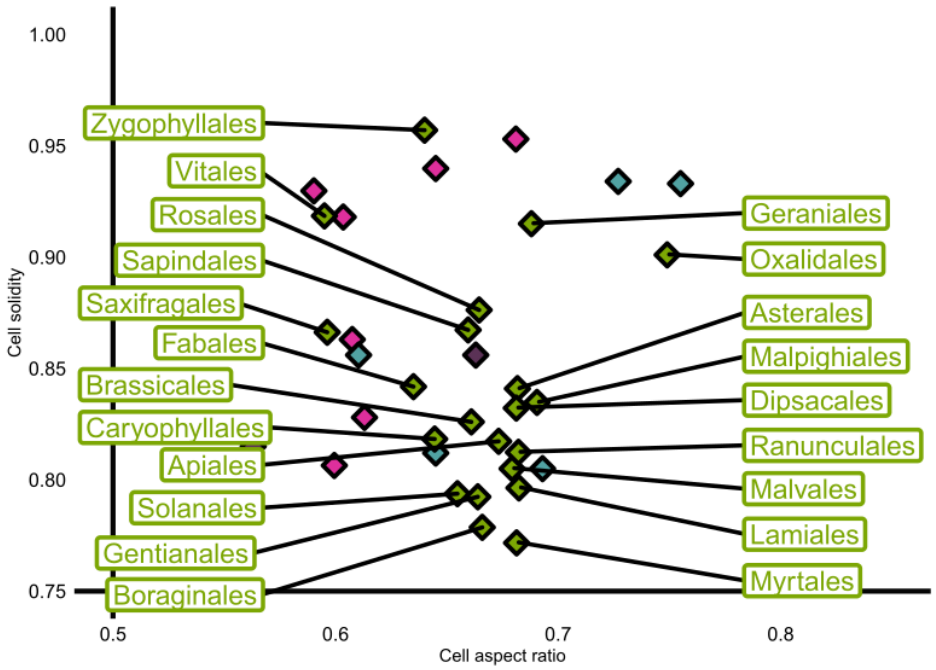
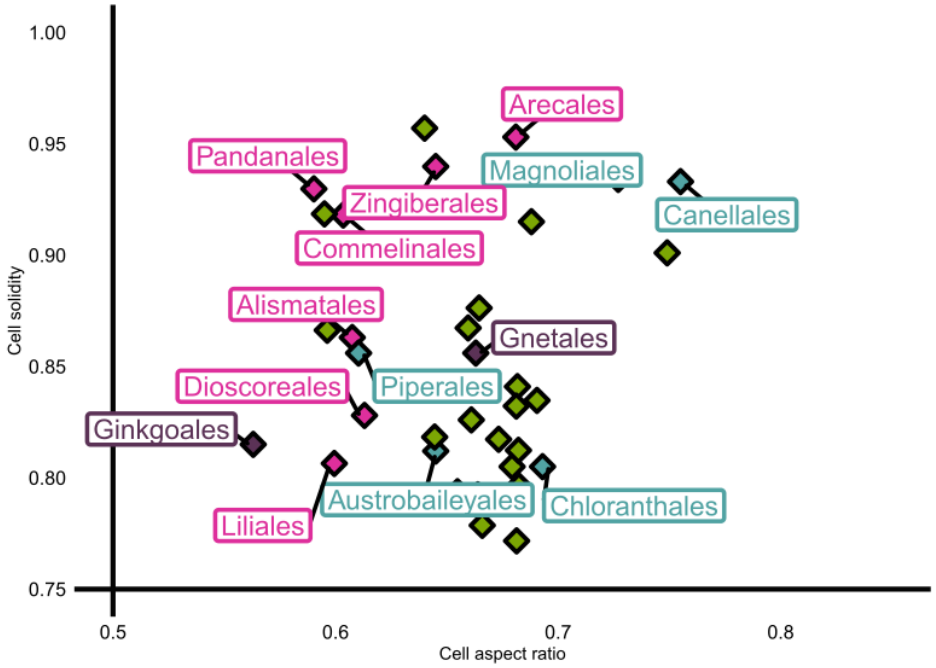


Figure 4.6 – Solidity against aspect ratio averaged by order
Cell shapes around the plot area illustrate aspect ratios of < 0.1, 0.2, 0.4, 0.6, 0.8 and 1 and solidity values of 0.6, 0.7, 0.8, 0.9 and 1. The colour of the data points mark the five groups, as shown in the legend. The two black lines correspond to an aspect ratio of 0.5 and solidity of 0.75 and divide the data into three groups. A close-up of the top right quadrant is presented in Figure 4.7.



(a) Orders with solidity > 0.75 and aspect ratio > 0.5 — Orders of the Eudicots are labelled



(b) Orders with solidity > 0.75 and aspect ratio > 0.5 — Orders of the other groups labelled

Figure 4.7 – **Solidity against aspect ratio averaged by order — solidity** > 0.75 **and aspect ratio** > 0.75 The full data is presented in Figure 4.6.

4.6 Phylogenetic correlations

The phylogeny of sampled taxa was extracted from a megaphylogeny of vascular plants [131] using phylomatic [132]. Where there was not an exact match for the species we sampled, we selected another species in the same genus from the megaphylogeny. The resulting phylogeny was used in tests for phylogenetic signal using the R package phylosignal [133]. The phylogeny figure was generated using ggtree [134], with final editing performed in Adobe Illustrator.

In qualitative surveys of epidermal cell shape, fern epidermal cells have been described as more undulating than eudicot epidermal cells, and, in turn, monocot epidermal cells as less undulating than eudicot epidermal cells [10, 126, 127]. To determine whether epidermal cell shape metrics were distinct to each clade in our dataset, and whether there were any trends in the evolution of epidermal cell shape, shape metrics were mapped onto a phylogeny of all the species sampled [131]. Cell solidity and cell aspect ratio for both the ad- and abaxial leaf faces were mapped onto the phylogeny. We tested for an underlying evolutionary signal in solidity or aspect ratio measures using statistical tests for phylogenetic signal [133].

Related species share a common ancestor and they may also share its phenotypic traits. As such, the presence of such a trait in related species cannot be treated as entirely random, independent occurrences. Revell *et al.* define the phylogenetic signal as the "statistical nonindependence among species trait values due to their phylogenetic relatedness" [135]. In more statistical terms, phylogenetic signal is defined as "the tendency for related species to resemble each other more than they resemble species drawn at random from the tree" by Blomberg and Garland [136]. The phylosignal package in R implements autocorrelation-based methods for tracking the phylogenetic signal in the data.

Moran's I index is often used as a global measure of autocorrelation in the dataset [137, 138, 139]. In order to establish the strength of the phylogenetic signal for single species, we used a localised version of this measure defined as

$$I_i = \frac{y_i - \bar{y}}{m_2} \sum_{j=1}^{n} w_{ij}(y_j - \bar{y}) \tag{4.1}$$

where y quantifies the trait examined, $\bar{y}$ is the mean, m_2 is defined by

$$m_2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n} \tag{4.2}$$

and w is a weight matrix based on the phylogenetic distance [140]. This is implemented by the function `lipaMoran` in the phylosignal package. The results are presented in Figure 4.8.

The ferns we sampled ($n = 2$) had strong phylogenetic signal for solidity (see Figure 4.8a), consistent with qualitative data in the literature [10, 13]. The cycads ($n = 7$) and some monocots ($n = 17$), including the grasses ($n = 7$) had a strong signal for cell aspect ratio with highly anisotropic cells (Figure 4.8b,c and d). In the case of the cycads, this was consistent with qualitative data in the literature [123]. Most core monocots, a monophyletic clade that includes all monocots apart from the Alismatales and Acorales, have cells with less complex cell margins, falling within the first two quartiles of solidity (values closer to 1; Figure 4.8c), consistent with the qualitative literature [10, 124]. While lower solidity values (higher undulation) were often present in the eudicots, the phylogenetic signal was weak: eudicot epidermal cells cannot be distinguished by solidity or aspect ratio values. The exception to this general trend was a subset of the Lamiales, where epidermal cells tend to have low solidity values (Figure 4.8g). In summary, while eudicot epidermal cell shape is not limited by any particular shape metrics, fern epidermal cells are characterized by high undulation depth, core monocot epidermal cells by low undulation, and many monocot and cycad epidermal cells by low aspect ratio.

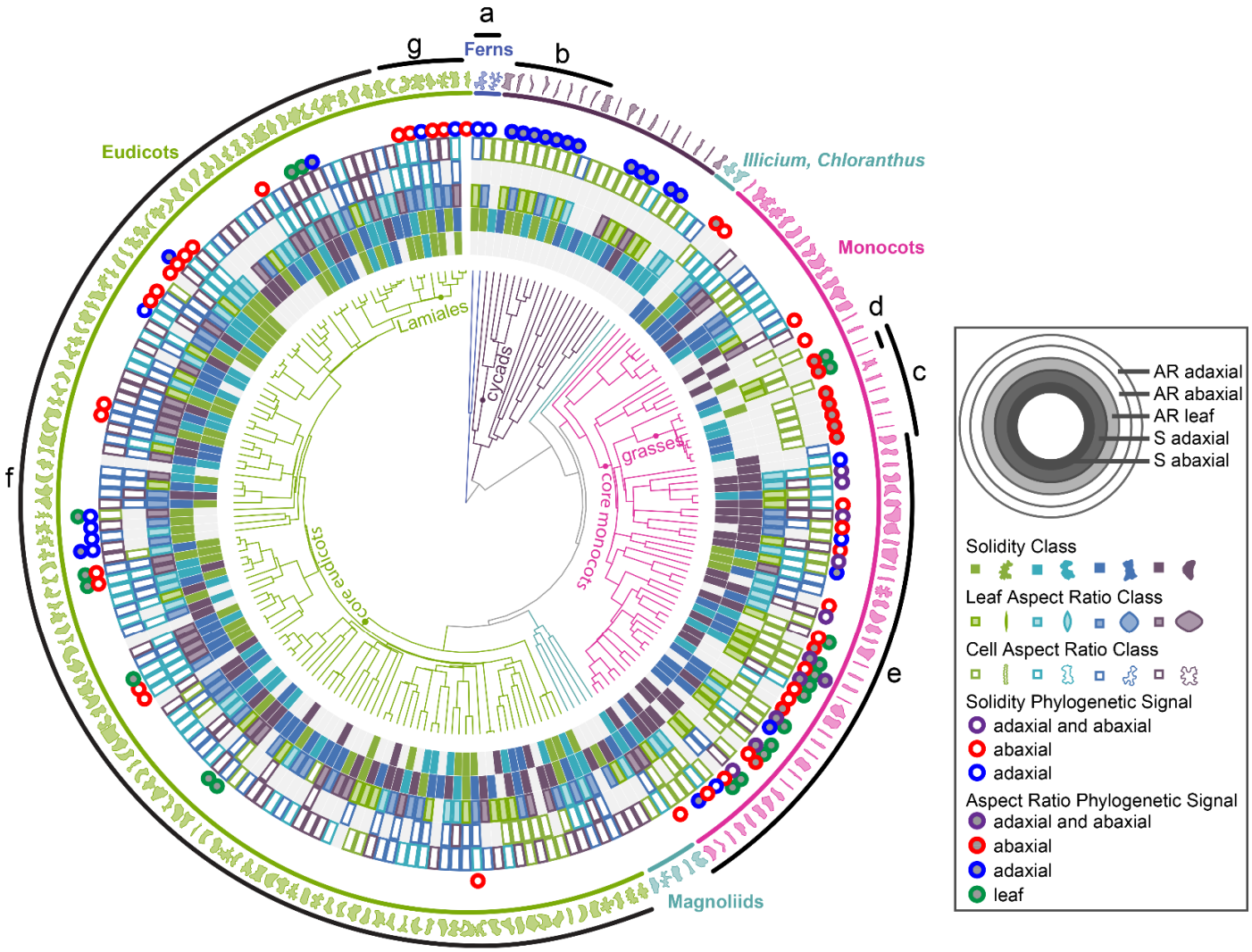


Figure 4.8 – Phylogenetic analyses of leaf and cell shape descriptors
Taxonomic groups are highlighted in colour, shape descriptors indicated in the rectangles. Circles mark the presence of a phylogenetic signal — see legend for positional key. Areas of particular interest are: a) ferns displaying a significant phylogenetic signal for cell solidity, b) cycads with a significant signal for cell aspect ratio, c) grasses with a significant signal for cell aspect ratio, d) grasses with a significant signal for leaf aspect ratio, e) core monocots showing a high occurrence of significant signal related to cell solidity, f) eudicots, where the phylogenetic signal is mostly weak, and g) a clade in the Lamiales with a significant signal for cell solidity.

4.7 Leaf shape

4.7.1 Geometric descriptors of leaves

Leaf shapes were also quantified using geometric descriptors: solidity and aspect ratio. Just like with cells, aspect ratio here describes the elongation of the leaf and solidity the quality of the outline. Solidity values lower than one can correspond to serrations on the margin or a lobed shape. This is presented on Figure 4.9.

On complex leaves, where a clear leaflet could be defined (in case of ferns for example), the shape of the leaflet was analysed. Leaf shapes were considered in order to compare them with the cell shapes observed and potentially reveal a relationship between cell shape and leaf shape. We reasoned that it is the "local" leaf shape that should correlate with the cell shape as the leaf shape, after all, is the sum of the cells in the leaf.

Low solidity points in monocots are curved leaves as shown in the graph — monocot leaves do not seem to develop a complex branching structure. Low solidity points in eudicots come from two species with a complex leaf shape and a somewhat random branching structure that could not easily be divided into leaflets and an elongated but curved leaf.

In fact, solidity does not appear to be a relevant quantity when looking at the leaf shape with the intent of comparing against the cell shape. Although the serrated outlines and complex leaflet structures are vital characteristics when identifying species, these features tell little about the growth of the cells in the middle portion of the leaf(let). Aspect ratio, on the other hand, reflects a global anisotropy that could prove more relevant on a cellular level.

4.7.2 Correlation of leaf shape and cell shapes

Cell shapes were also represented by solidity and aspect ratio. Correlations between the two leaf and two cell shape characteristics were examined separately for monocots and eudicots, the two largest groups in the sample set.

Cell aspect ratio — leaf aspect ratio

Figure 4.10 shows the average cell aspect ratio versus the leaf aspect ratio. Example cell and leaf shapes are inserted in the panel of Eudicots corresponding to aspect ratios of < 0.1, 0.2, 0.4, 0.6, 0.8 and

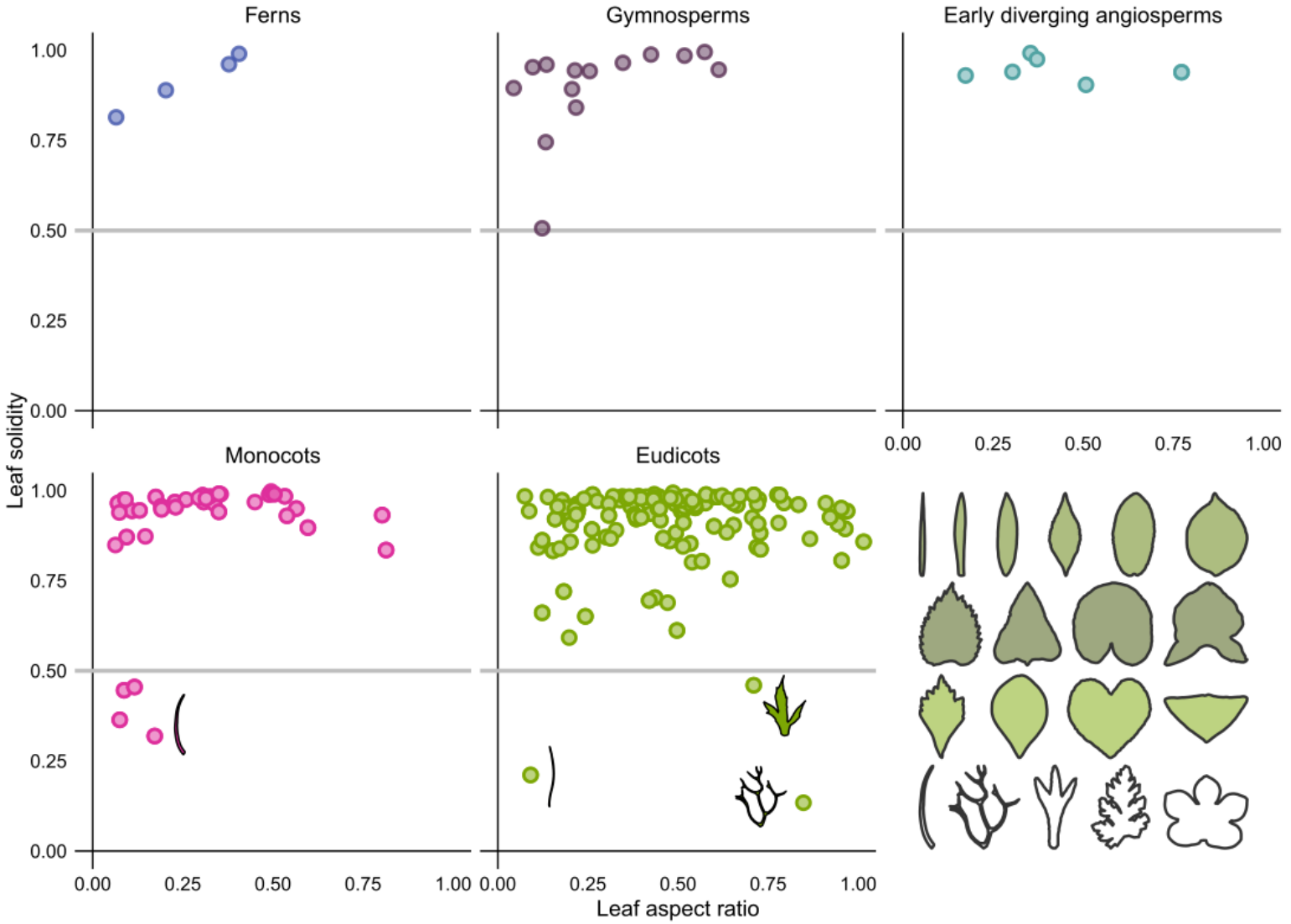


Figure 4.9 – **Leaf solidity and aspect ratio** Panels 1-5 show the values for the five groups, panel 6 shows example leaf outlines: row 1 — varying elongation, row 2 — skewed towards the bottom, row 3 — skewed towards the top, row 4 — low solidity shapes. Grey line corresponds to a solidity value of 0.5

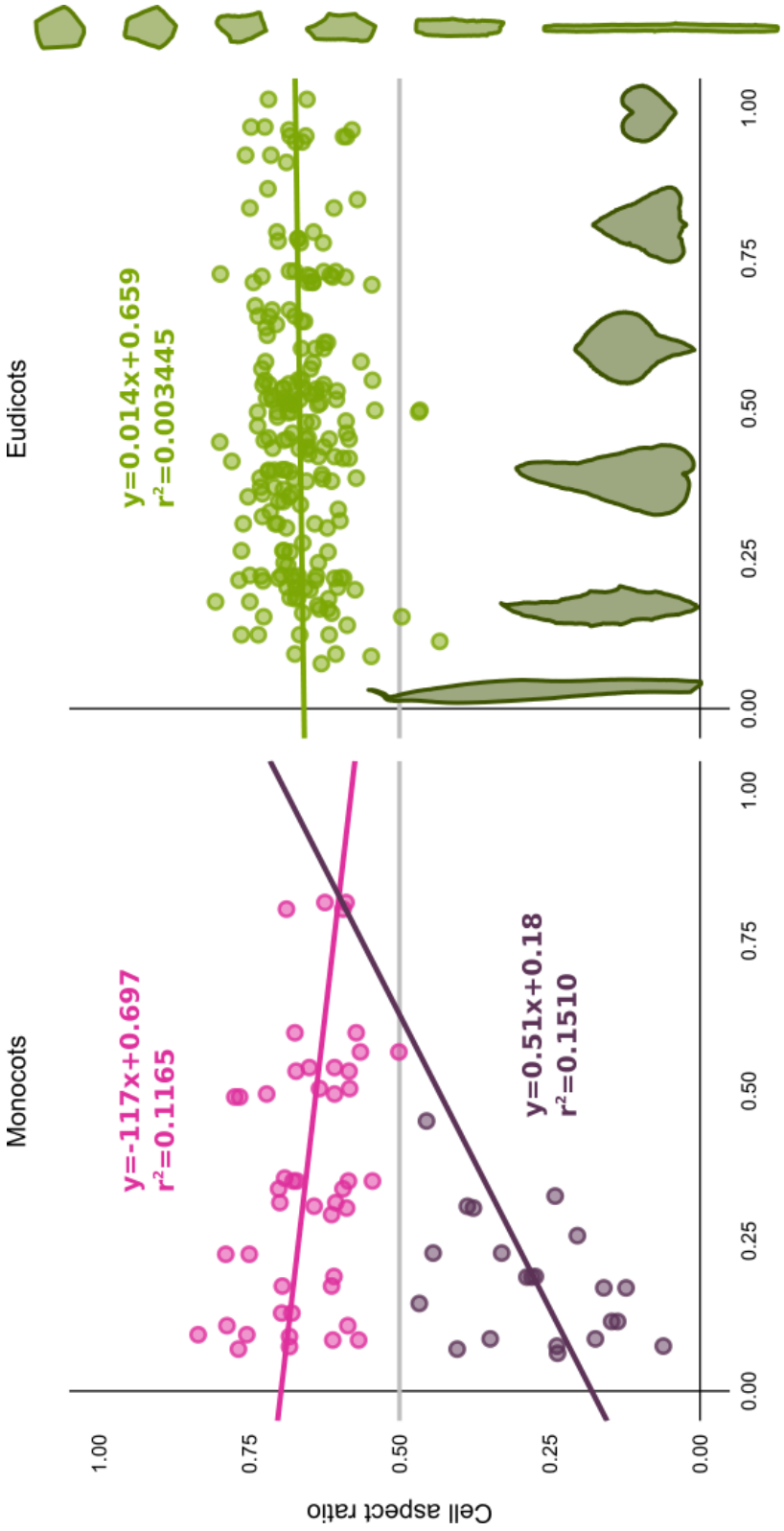


Figure 4.10 – Correlations between cell and leaf aspect ratio — data presented for Monocots (left) and Eudicots(right). Example leaf and cell shapes correspond to aspect ratios of < 0.1, 0.2, 0.4, 0.6, 0.8 and 1. Monocots are divided into two groups according to cell aspect ratio. Pink marks species with cell aspect ratio greater than 0.5, violet marks species with cell aspect ratio smaller than 0.5.

1. Leaf aspect ratio covers most of the available range, down to 0.08, corresponding to a 1:12 ratio of width to length. Cell aspect ratio of monocots also covers nearly the full range available while eudicots mostly stay above $0.3 - 0.4$, corresponding to a ratio of approximately 1:3.

While eudicots are evely spread out in the accessible range, showing no correlation between cell elongation and leaf elongation (Spearman's $\rho = 0.0136$), monocots can be divided into two groups: one that occupies similar coordinates as the eudicots and does not correlate with the leaf elongation at all ($\rho = -0.1825$) and another (highlighted in purple) consisting of cells more elongated than eudicots, which shows a weak correlation with the leaf aspect ratio ($\rho = 0.3650$). This indicates that extremely anisotropic cell expansion results in extremely anisotropic leaf expansion. Extremely elongated leaves, however, can also be formed by less elongated cells.

Cell solidity — leaf aspect ratio

Cell solidity with the monocots stays typically above 0.7, with eudicots it can reach 0.5 and beyond. Correlations with cell solidity are very weak: Spearman's ρ for the full sample set is -0.0906, -0.1680 for monocots and 0.0323 for eudicots.

Cell aspect ratio — leaf solidity

Leaf solidity — with a few exceptions — is confined to 0.8 in monocots and 0.6 in eudicots. Leaves with very low solidity in the monocots are, without exception, very long and slightly curved leaf. Leaves with very low solidity in the eudicot are not an artefact: these have complex shapes.

Based on this data it seems that there is little correlation between the complexity of the leaf and the aspect ratio of cells. Spearman's ρ for the full dataset is 0.0336, 0.1123 for monocots and 0.0846 for eudicots, suggesting that the complexity of the leaf outline and the elongation of cells are independent of each other.

Cell solidity — leaf solidity

The density of points in the top left corner of both graphs suggests that most cells and most leaves have, in fact, simpler shapes. Spear-

man's ρ here takes 0.1477 for the full dataset, indicating a slightly stronger relationship than between cell aspect ratio and leaf solidity. For monocots, $\rho = 0.1765$ and for eudicots, $\rho = 0.0782$.

These values can be slightly higher than those from the cell aspect ratio–leaf solidity comparison because cell solidity covers a smaller range and the data points are more confined, also, the density of points is very high around the top right corner. Based on this graph, there is no strong trend between cell solidity and leaf solidity.

4.8 Abaxial and adaxial sides

Another pattern mentioned in the qualitative literature is different cell shapes abaxially vs. adaxially: abaxial cells may have more undulate margins [13]. In our dataset we had adaxial and abaxial data for 111 species (81 eudicots and 30 monocots). To examine whether abaxial cells had more undulation, we calculated the average solidity between the two sides of the leaf and the difference between the average adaxial solidity and the average abaxial solidity. This is shown on Figure 4.11.

For the eudicots, the solidity difference between sides covered a broader range, indicating that species ranged from a greater to lesser difference between sides. Within the monocots, the difference between the sides was less pronounced and the average solidity was closer to 1, indicating a less undulating margin as compared to eudicots (Figure 4.11; monocots, pink).

This result was consistent with our analyses presented in the previous sections. In both monocots and eudicots, the majority of the difference data was positive, indicating that when a difference in cell solidity is present, the abaxial cells tend to have more undulations (Figures 4.8 and 4.11). This is in line with the qualitative literature [10, 125].

The cause of such differences are ripe for discovery. In many cases, different sides of the leaves experience different microclimates; undulation exhibits some environmental plasticity and thus it is plausible that more undulation on abaxial surfaces could relate to local environmental influences [23, 21, 12]. In addition, abaxial vs. adaxial developmental identity may contribute to differential undulation [23, 15, 19]. Lastly, it is possible that differential growth rates between the two sides may relate to differential undulation; however, these differences would need to be temporary to finally

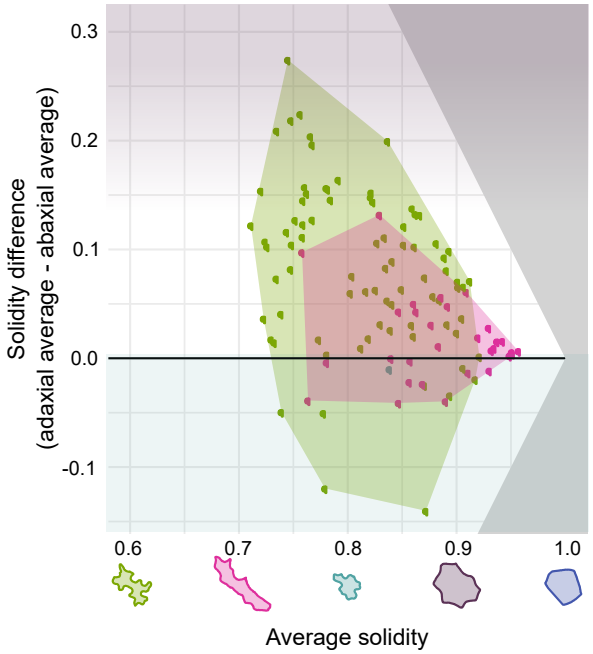


Figure 4.11 – Comparing abaxial and adaxial undulations
Horizontal axis shows the mean solidity of the two sides, vertical axis shows the adaxial-abaxial difference. Grey areas mark the plot area inaccessible for cells. Samples with a positive difference have more undulating abaxial cells, samples with a negative difference develop more undulations in the adaxial epidermis.

yield a flat leaf.

The difference in area and aspect ratio between cells from ad- and abaxial sides was also examined. It was found that 75 species had an average adaxial area at least 5% greater than the average abaxial area and 19 species had an abaxial area at least 5% greater than the adaxial area. In terms of aspect ratio, 25 species had an average adaxial aspect ratio at least 5% greater than the abaxial average and 17 species had an abaxial average at least 5% greater than the average. These distributions are a lot less significantly skewed towards the adaxial side.

4.9 Discussion

From the comparison of the two methods, geometric descriptors were found to work better in quantifying undulations in the outline, with the help of solidity. But because solidity is a global measure, it cannot distinguish between cells with different lobe structures: a cell with a round base shape

and moderately deep undulations on the outline typically has similar solidity values to a complex shape with a few deep lobes. Being able to measure this difference could lead to a better classification, therefore it is of interest to try to find descriptors that capture the structure. Most of the cells, however, do not have lobes: a method focusing on lobes would therefore be inappropriate. The persistent homology tool described by Li *et al.* [106] seemed the most promising as it is capable of capturing detailed topological information and an attempt was made to use this tool to further elucidate the differences in cell shape.

The exciting details, however, come at a high computational cost. The calculations take a long time and require a lot of resources. The resolution with which we can define the cell outline at the input is limited by the available RAM on the computer. On an average computer with 4 GB RAM, the maximum number of vortices defined in the outline with which the calculations could still safely run was found to be 500. 500 vortices is close to the limit where features from certain cells are starting to get lost, but in general it is sufficient to capture features of the outline.

Initially, the time for completing the calculations for one sample was around 80–100 s, making the total time required for completing the calculations a little over 10 days. However, over the course of several days, the speed of computations dropped significantly, and during a week-long period of undisturbed run only about 1200 cells were analysed, making the total runtime estimate for the 10260 cells in the dataset over 8 weeks.

Naturally, the time demand can be reduced by running the code on a more powerful computer or reducing the number of points kept on the outline and losing part of the data. However, this raises the question of what makes a method good: given the time difference between the geometric method and the persistent homology one, the loss of structural information is a price I was willing to pay.

4.10 Conclusions

In this experiment we collected and examined the epidermal cell shapes of 230 species of vascular plants. These cells show a broad variation in size and shape, most importantly in elongation and the quality of the outline. These differences were successfully quantified using geometric descriptors and were

used to examine phylogenetic correlations, correlations with the leaf shape and the differences between the abaxial and the adaxial sides.

Elliptical Fourier descriptors were found to perform poorly on this data set as outline undulations were present at many different frequencies. While geometric descriptors do not retain structural information about the lobes, that may be important for the most complex cell shapes, they are quick to compute and easier to interpret.

Strongly undulating outlines, like those of *Arabidopsis* were found to be uncommon amongst vascular plants, with ferns being the only exception: the four species we examined (arguably a low number) did have some of the most complex cell shapes. Eudicots are the most diverse, with many species displaying some level of lobing in their cell outlines and covering a wide range in solidity. The variation itself is different in different families. In some families, most species are centred on similar solidity values, in others, even species from the same genus can have large differences in solidity. The variation within a species is interesting: this relates to the way cells are organised in the epidermis. Some families contain species that have undulating outlines at a nearly constant solidity, while others have cells varying strongly in solidity. The difference in ab- and adaxial undulations is another characteristic that varies from family to family.

Chapter 5

The tales of *thaliana*

5.1 Motivation

Arabidopsis thaliana is, without doubt, the most important model organism in plant science. Discovered in 1577 by the German botanist Johann Thal (Ioannes Thalius), this small plant entered the world of botany under its maiden name, *Pilosella siliquosa* [141] and was later renamed *Arabidopsis thaliana*, in honour of its discoverer.

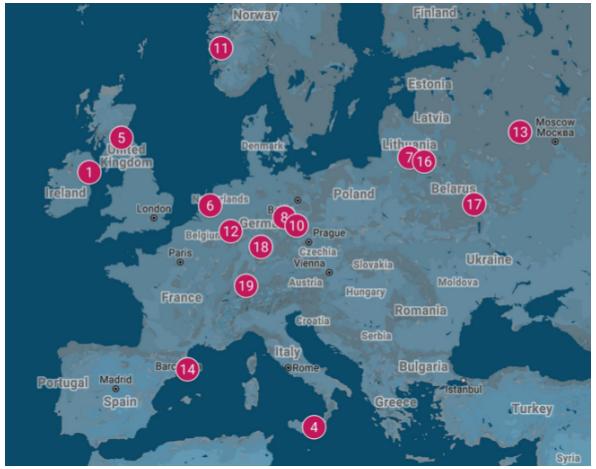
Being quite tiny, *A. thaliana* does not require much space to grow and has a rapid life cycle of about 6 weeks. Most often, it reproduces through self-fertilisation and can produce tens of thousands of seeds per plant [142] — these properties make *A. thaliana* very convenient to work with. Its genome is quite small compared to other flowering plants at only 125 megabases on five chromosomes, and it was the first one to be fully sequenced in 2000 [143]. Having this information available is what made *A. thaliana* truly the most important plant for a wide range of studies. Its popularity lead to the creation of a uniquely rich spectrum of tools and resources.

One such tool developed was the set of Multiparent Advanced Generation Inter-Cross (MAGIC) lines [144]. They are descendants of 19 accession lines that were selected based on their location of origin to cover a large area, or because they are very frequently used in studies. These lines were created to allow a higher resolution mapping of quantitative traits: genetic data of the parent lines along with the resulting inter-cross (MAGIC) lines was made available. Figure 5.1 shows the available data about the locations where the nineteen MAGIC parent lines were originally collected.

number	line	latitude	longitude	max. altitude
1	Bur-0	54.1	-6.2	100 m
2	Can-0	29.2144	-13.4811	1260 m
3	Col-0	38.3	-92.3	100 m
4	Ct-1	37.3	15.0	100 m
5	Edi-1	55.9681	-3.21833	100 m
6	Hi-0	52.0	5.0	100 m
7	Kn-0	54.89852	23.90359	100 m
8	Ler-0	47.984	10.8719	600 m
9	Mt-0	32.34	22.46	200 m
10	No-0	51.0581	13.2995	300 m
11	Oy-0	60.385543	6.193019	100 m
12	Po-0	50.72138	7.08981	100 m
13	Rsch-4	56.25839	34.32834	200 m
14	Sf-2	41.7833	3.03333	100 m
15	Tsu-0	34.43	136.31	100 m
16	Wil-2	55.0	25.0	200 m
17	Ws-0	52.3	30.0	200 m
18	Wu-0	49.7878	9.9361	200 m
19	Zu-0	47.3667	8.55	600 m

(a) Geographic data

(b) MAGIC parent lines across the world



(c) MAGIC parent lines from Europe

Figure 5.1 – Geographic data of the MAGIC parent lines
Data from The Arabidopsis Information Resource [145], missing points for the altitude estimated from the latitude and longitude coordinates. Map prepared with Google Maps.

Along with the many natural variants of *A. thaliana*, there are also many mutants available that have been used to study the function of particular genes — including those dedicated to unveil the biochemistry of lobe growth. As presented in Section 1.5 of the introduction, for a lobe to grow there are two main factors needed: anticlinal microtubule bundles restricting the growth at the necks and a fine cortical actin network at the tips of the lobes. The key factors affecting the formation of microtubule bundles or the actin network have been identified: ROP6 and RIC1 are responsible for microtubule organisation, ROP2 and RIC4 promote the formation of the actin network. Mutations in these proteins lead to a loss of function that leads to differences in the cell shape.

Recent studies quantify the differences in cell shapes between wild type and mutant cells showing pavement cell shape defects using traditional morphometric measures such as lobe width and lobe length [68, 66]. While this is very useful information, it is hard to automatise in a reliable fashion. More novel methods can find lobes automatically but are not able to characterise them [105]. The drawback of lobe-based methods is that this quantification, while suitable for this species, cannot be used to compare against other species as the majority of species do not grow pavement cells with such intricate lobe structures — demonstrated in Chapter 4. As *A. thaliana* is used as a model organism representing other flowering plant species too, it is of interest to test a method that allows a comparison with other species.

This chapter presents a study of the natural variation in *A. thaliana* using the method developed in the previous chapters.

5.2 Aims

The aim of this project was to examine whether there are subtle variations in epidermal cell shapes within a species, and how well we can quantify it using our previous methods. The chosen species was *Arabidopsis thaliana*, and the chosen lines were the nineteen parent lines used to create the MAGIC lines [144]. The MAGIC parent lines were selected not because of previous evidence that they show a great variation in shape but because they offer the possibility to relate any significant differences to genetic data.

To provide a scheme for normalising variations in the cell shapes between different lines, a handful of mutant lines showing or expected to show cell

shape defects was also selected. Mutants showing defective cell shapes are expected to differ more from wild-type plants than the natural variation amongst accession lines. The mutant lines selected were *arp3*, where the mutation affect the functioning of the ARP2/3 complex and known to have lobeless cell shapes [47]; *rop2* and *ric4*, where the mutation occurs in the ROP2 and RIC4 proteins also affecting the actin network [43, 39]. The mutants *rop6* and *ric1* carry mutations in the ROP6 and RIC1 proteins respectively and both result in cell shapes where the microtubule bundle formation is affected, leading again to a moderate change in cell shapes [68]. At last, *qua1* was also selected as it was found to have serious cell shape defects due to adhesion problems, but with the factors responsible for the formation of the actin network and the microtubule bundles necessary for lobe growth undisturbed [146].

Experimental methods are detailed in Section 5.3. Geometric descriptors were used for the major part of the analysis. Individual results by groups are presented in Section 5.4, correlations with the geographic data are examined in 5.5 and principal component analysis based on the geometric descriptors is discussed in Section 5.6. Results are discussed in Section 5.7 and conclusions are drawn in 5.8.

5.3 Experimental methods

5.3.1 Plant material

Sterilised seeds (10 from each of the lines studied) were planted on 0.5 MS plates and, following cold treatment at 4 °C for at least 48 h, grown under 16 h light/8 hour dark cycles for 7 days, at a temperature of 20 °C. Seeds from the MAGIC parent lines were kindly provided by Katie Abley, and seeds from the mutants studied in this experiment originate from Firas Bou Daher. Geographic data originates from The Arabidopsis Information Resource (TAIR) [145].

5.3.2 Staining and imaging

For every line, three cotyledons were cut from three different seedlings, placed on a glass slide and stained with a freshly made $0.02\,\mathrm{g\,l^{-1}}$ propidium iodide solution. After a few minutes, samples were covered with a cover

slip and imaged in a Leica TCS SP8 confocal microscope with a 10× non-immersion objective. Excitation beam was provided by a diode-pumped solid state laser at 561 nm and wavelengths between nm and nm were collected by a HyD detector. Laser power was adjusted as necessary and typically kept between 30 and 40% of the full power, 20 mW.

In order to fit a number of cells into the viewing field and not lose visibility of the outline features, display resolution was raised from the standard 512×512 to 2048×2048, corresponding to an actual resolution of 0.57 µm per pixel at zoom = 1.00. Zoom was kept at either 1.00 or 0.75×. To further improve image quality, final images were recorded with line averaging (4 lines). When necessary, due to the curvature of the cotyledon for example, stacks were taken instead of single images. In these cases, a maximum projection of these stacks was also saved and used for segmentation.

5.3.3 Segmentation

The images obtained this way were suitable for automatic segmentation with minimal human interaction. Images were segmented using CellSet [117]. The result of this segmentation was then corrected by hand. From each cotyledon imaged, 30 cells were selected for analysis, resulting in a dataset of 1710 cells from the MAGIC parent lines and 180 cells from mutant lines.

5.3.4 Shape analysis

As previously, outlines from the segmented images were extracted in MATLAB [118], and the list of x–y coordinates for every shape was further analysed using the Momocs package in R [119]. Shapes were first studied using geometric descriptors in a separate manner, then subjected to PCA. Fourier analysis was also attempted and gave poor results.

5.4 Cell shapes by individual descriptors

Cell outlines were quantified using four descriptors: area, aspect ratio, circularity and solidity, as defined in Chapter 2. Mean values for each line studied are presented on Figure 5.2. The distribution of the values of each descriptor is shown for every sample in Appendix D.

Figure 5.2 summarises this data for the MAGIC parent lines and the mutant lines studied. Median values of aspect ratio for each line is shown against median circularity. The fill colour of the marker corresponds to the magnitude of the difference between the median area for the given line and the median area for all of the wild type lines (red marks a positive difference — larger cells, blue marks a negative difference — smaller cells). The deviation is expressed as percentage of the wild type median.

As the variations are typically non-Gaussian, the Mann-Whitney test was used for determining whether there are significant differences between two groups in the data and Kruskal-Wallis test was used on more than two groups to test for significant differences. As per general practice, differences resulting in p values smaller than 0.05 were declared significant. These rank-based tests typically pick up variations in the median. These tests were used to identify statistically significant differences between MAGIC parent lines.

Figure 5.2 also shows the classification obtained on this basis. Classification based on circularity is denoted by the marker shape used for the particular group and the colour of the label. Classification based on area is marked by the outline colour of the marker. Mutant lines were not included in these classification, these were only compared to wild type lines later. There were no significant differences in aspect ratio between the MAGIC parent lines.

Area

Clear outliers in terms of area are Ler-0 and Sf-2: these are significantly different from any other groups. The area of pavement cells in the case of Ler-0 is significantly smaller, while in case of Sf-2, significantly higher than those of the other lines. The rest of the wild-type lines form two groups, within which the differences are not significant: group 1 consists of Bur-0, Edi-1, Hi-0, Tsu-0, Wil-2, Ws-0, Wu-0 and Zu-0 (higher median values); group 2 includes Can-0, Col-0, Ct-1, Kn-0, Mt-0, No-0, Oy-0, Po-0 and Rsch-1 (lower median values).

Mutants — except for *arp3* — typically have smaller cell areas and show smaller variations. Correlation between the mutant line and the full wild type sample set show that the *rop2*, *rop6* and *ric4* lines are not significantly different from wild type cell shapes, unlike those of the *qua1* and *ric1* lines.

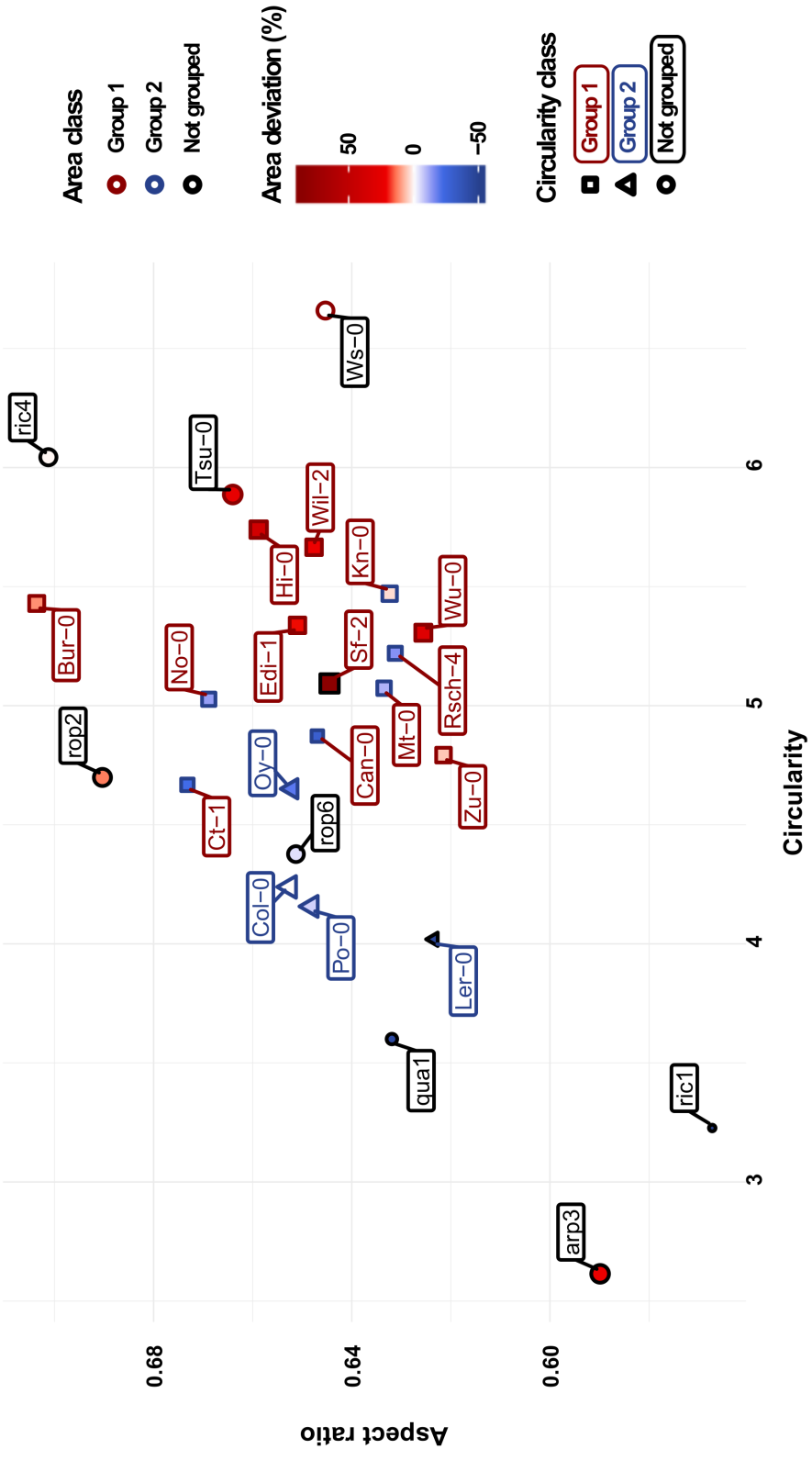


Figure 5.2 – Geometric descriptors of MAGIC parent and mutant lines
Median values of aspect ratio is shown against circularity for each line. Differences in aspect ratio between different MAGIC parent lines are not statistically significant. MAGIC parent lines fall into two distinct classes based on circularity except for Tsu-0 and Ws-0. Marker shape and label colour indicate circularity classes as shown in the legend. The fill colour corresponds to the difference between the median area of the group and the median area of wild type cells, expressed as percentage of the wild type median. MAGIC parent lines can also be grouped based on area, except Ler-0 and Sf-2. This is indicated by the outline colour of the marker.

Aspect ratio With the general sample set presented in Chapter 4 and with the developing maize cells described in Chapter 3, aspect ratio the most important descriptor out of the four geometric descriptors and the main characteristic of the first principal component arising from the Fourier analysis. In the case of *Arabidopsis*, however, aspect ratio is of little importance: the Kruskal-Wallis test shows that no group is significantly different from the others.

This quantity, despite being reasonably similar between different wild type lines, shows a clear difference with some of the mutants. In particular, correlations between the full wild type dataset and each mutant lines show that *arp3* and *ric4* cells have significantly different aspect ratios: those of *arp3* are significantly lower, indicating that these cells are more elongated. In contrast, *ric4* has less elongated cells than the average wild type cell. No significant difference exists in the case of *qua1* and *rop6* in aspect ratio and *ric1* and *rop2* are on the boundary with Mann-Whitney p being 0.05168 and 0.07822 respectively.

Circularity It was shown in Chapter 2 that circularity is not always an ideal measure as it varies strongly with elongation (aspect ratio). Since aspect ratio is not a distinguishing factor amongst the MAGIC parent lines, circularity here is a good measure for quantifying shape. In fact, as it employs the perimeter to quantify the complexity of the cell shape, it captures more detailed information about the lobe structure, if only globally, than solidity, which makes it a more suitable measure for *Arabidopsis*.

There are two wild type lines that show significant differences in circularity: Tsu-0 and Ws-0. The remaining lines can be divided into two groups: Bur-0, Can-0, Ct-1, Edi-1, Hi-0, Kn-0, Mt-0, No-0, Rsch-4, Sf-2, Wil-2, Wu-0 and Zu-0 form the first, and Col-0, Ler-0, Oy-0 and Po-0 the second.

Most mutant lines again show significant differences: *arp3*, *qua1* and *ric1* all have p values well below 0.05, while *ric4* and *rop6* have p values around 0.09. The only mutant line not having significantly different circularities is *rop2*.

Solidity When the base shape of cells is similar, solidity cannot reliably distinguish between cells with lots of shallow lobes or a few deeper ones. Solidity values are not shown on Figure 5.2 to avoid redundancy, but the distribution of solidity values can be found in Appendix D. Nevertheless, it is still a good measure for quantifying the global amount of deviations from a convex shape.

The results are quite similar to those of circularity in qualitative terms. Again, Tsu-0 and Ws-0 were found to be significantly different from any of the other groups. Most of the remaining lines fell into the same groups as for circularity, with the exception of Ct-1, that is more similar to Col-0, Oy-0 and Po-0 in terms of solidity.

Regarding mutant lines, *arp3*, *qua1* and *ric1* were again found to be significantly different from the wild type samples, but *ric4* in this respect is not significantly different from the wild type samples ($p = 0.77$) along with *rop2*. Still close to the boundary with $p = 0.14$, *rop6* was still not found significantly different.

In summary, examining these four variables we found that some MAGIC parent lines differ from the others in area and shape complexity (circularity and solidity) but not in aspect ratio. The differences of the cell shapes in the mutants studied was subject to such limitations: some mutant lines were significantly different in aspect ratio as well. Table 5.1a presents the median values of the four descriptors and the groups determined on the basis of significant differences for the MAGIC parent lines. Table 5.1b presents the median values for the mutant lines along with the p values compared to the entire wild type dataset.

line	area (µm^2)		aspect ratio	circularity		solidity	
Bur-0	6077	1	0.703	4.92	1	0.635	1
Can-0	3817	2	0.647	4.49	1	0.631	1
Col-0	5455	2	0.650	3.92	2	0.672	2
Ct-1	4267	2	0.684	4.30	1	0.649	2
Edi-1	6571	1	0.650	4.83	1	0.624	1
Hi-0	7637	1	0.663	5.20	1	0.612	1
Kn-0	5640	2	0.630	4.95	1	0.626	1
Ler-0	2920	*	0.623	3.81	2	0.646	2
Mt-0	4898	2	0.633	4.63	1	0.631	1
No-0	4750	2	0.669	4.58	1	0.621	1
Oy-0	4451	2	0.655	4.26	2	0.652	2
Po-0	5120	2	0.654	3.83	2	0.670	2
Rsch-1	4652	2	0.631	4.77	1	0.634	1
Sf-2	10239	*	0.644	4.61	1	0.641	1
Tsu-0	6799	1	0.662	5.43	*	0.599	*
Wil-2	6636	1	0.648	5.15	1	0.612	1
Ws-0	5439	1	0.642	6.06	*	0.584	*
Wu-0	7118	1	0.628	4.93	1	0.619	1
Zu-0	5759	1	0.622	4.40	1	0.630	1
all wild type	5486		0.646	4.70		0.631	

(a) Median values of the four geometric descriptors by MAGIC parent line

Integers next to area, circularity and solidity values mark the group the given line falls in, stars denote lines that are significantly different from all of the others based on the given quantity.

line	area (µm^2)	p	aspect ratio	p	circularity	p	solidity	p
arp3	6698	0.098	0.587	0.035	2.45	0.000	0.741	0.000
qua1	3031	0.000	0.631	0.867	3.35	0.001	0.684	0.002
ric1	2620	0.001	0.565	0.052	3.03	0.000	0.716	0.000
ric4	5489	0.940	0.702	0.013	5.53	0.090	0.606	0.071
rop2	6172	0.828	0.690	0.078	4.30	0.715	0.624	0.977
rop6	5228	0.554	0.650	0.896	4.00	0.095	0.652	0.143

(b) Median values of the four geometric descriptors for mutant lines and Mann-Whitney p values compared against the full wild type dataset

Table 5.1 – Median values of geometric descriptors for MAGIC parent and mutant lines

5.5 Correlation with geographic parameters

The nineteen MAGIC parent lines originate from different locations on Earth. Although most of them come from Europe, Col-0 arrived from North America, Tsu-0 from Asia and Can-2 and Mt-1 from Africa. These locations spread over a wide range of latitude and longitude and also include different altitudes. It is known that environmental conditions such as the amount of sunlight available or the humidity can influence the cell shapes of plants [13] and therefore it is possible that different cell patterning is prioritised under different climates, to such an extent that it exerts a selection pressure. If this were true, we would expect to see a correlation between cell shape features and climate data.

The availability of climate data is, however, scarce: TAIR only provides this data for a few of the MAGIC parent lines. In lieu of climate data, geographic parameters were used to examine correlations between cell shape and conditions at the original locations. Spearman's ρ was used to determine whether there are any monotonic tendencies present in the data. These values are presented in Table 5.2 below. Spearman's test tells us whether there is a monotonic relationship between the two quantities. As it is less reliable when there are many y values corresponding to one x, the test was carried out on the median for each line.

	Area	Aspect ratio	Circularity	Solidity
Latitude	0.0965	0.0544	0.2561	-0.1308
Longitude	-0.0807	-0.2326	0.4298	-0.4723
Altitude	-0.4056	-0.5497	-0.1295	-0.2347

Table 5.2 – Correlations between geographic coordinates and shape measures, Spearman's ρ

According to the data, there are no strict monotonous relationships between geographic parameters and shape descriptors. The strongest correlation is between altitude and aspect ratio, however the correlation coefficient stays below 0.7 that is usually accepted in biological studies. Another interesting observation is the relatively strong correlation of circularity and solidity with longitude. This trend is also very weak and may entirely be influenced by the two points at the extremes of the longitude scale: Col-0, showing relatively less deeply lobed cells, and Tsu-0, showing more strongly lobed cells.

This analysis failed to measure a significant correlation between location and cell shape. However, the relationship between geographic coordinates and climate is not straightforward: the lack of correlation with geographic coordinates does not therefore exclude the possibility of a relationship between climate variables and cell shapes.

5.6 Combining geometric descriptors

Following the individual analysis, the four geometric descriptors were combined and subjected to principal component analysis as before, with the unbounded variables — area and circularity — logarithmised. Figure 5.3 presents the cell shape data grouped by lines on the PC1–PC2 morphospace, along with the directions representing the original geometric descriptors in the plane.

The first two principal components were found to capture 65.1% and 24.7% of the variability in the data respectively, which provides a good description of the raw geometric descriptor dataset. PC1 is composed mostly of area, circularity and solidity, while PC2 is almost entirely capturing aspect ratio. Example cells aid in understanding the meaning of the principal components: as PC1 decreases, the cells become larger and the cell shape develops more and more complex lobes. As PC2 increases, the cells become more elongated.

This is a novel behaviour: in the previous study (Chapter 4) area varied with the aspect ratio, while here area varies with the outline parameters solidity and circularity. With increasing area, circularity increases and solidity decreases: the bigger the cell grows, the more complicated its outline becomes. It is important to highlight that circularity and aspect ratio are dimensionless quantities and as such, independent of scaling. Therefore, this variation between area and shape parameters really does reflect the manner of growth.

Looking at the individual panels for different lines, we see that the point clouds spread out and cover a wide range of the morphospace, especially compared to the separation of the groups. The only variation visible from the plot is that while most lines seem to cover the same range, a few concentrate into smaller areas. For example, Col-0 is fairly uniformly spread out, signifying that its cells take very diverse shapes and sizes, while Tsu-0

clusters around the left side of the plot, indicating that the sample contains predominantly bigger cells with a more complex shape.

Actin mutants show some deviations on this plot: *arp3*, although quite spread out, is shifted to the right, marking the presence of smaller and less complex cells. In addition, as the comparatively high PC2 values suggest, these cells are more elongated. The mutant *rop2*, in contrast, is shifted towards the bottom left corner, indicating less elongated cells (PC2) and cells of greater area and/or complexity (PC1). From this measure it is not obvious whether the location on the PC1 axis is due to the area or the shape complexity. At last, *ric4* covers a similar range in PC1 to most wild type cells, but PC2 values show that these cells are on average less elongated.

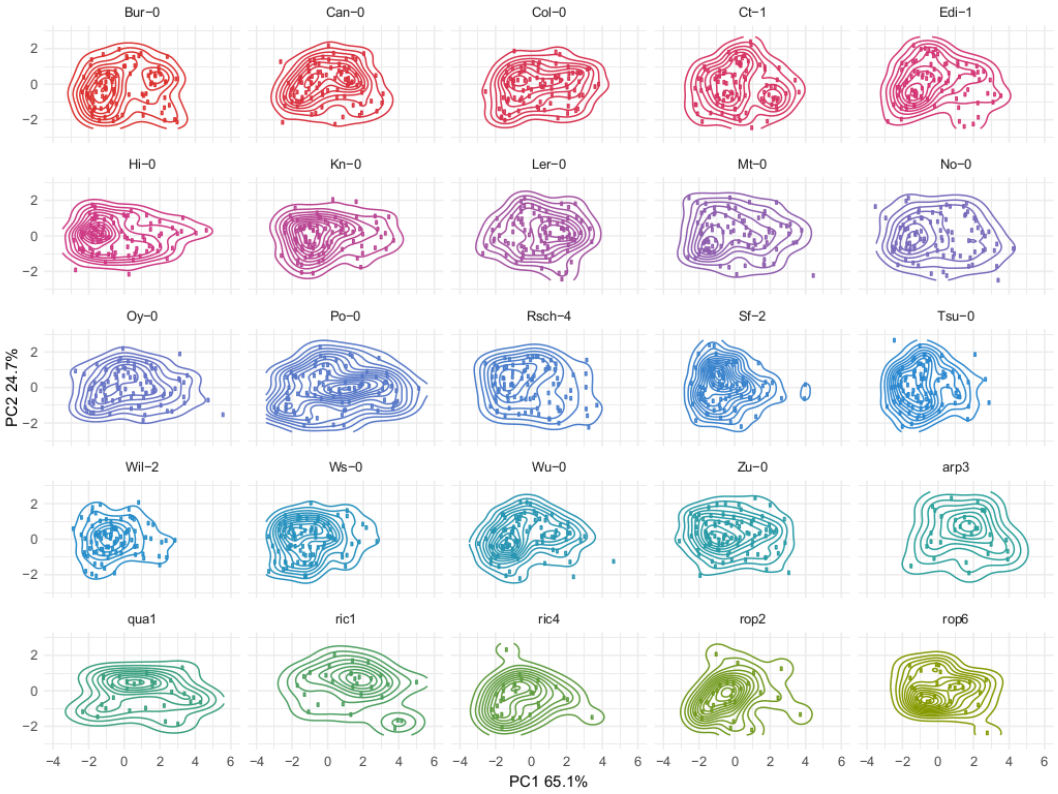
Disruptions in the formation of microtubule bundles also appear in this analysis: *rop6* occupies a central region, with most of the data falling in the bottom left quadrant, indicating that the cells in this sample are slightly complex and less elongated than in *arp3*, but smaller and less complex than in wild type seedlings. Meanwhile, *ric1* occupies an area similarly spread out to those of the wild type cells but limited to higher PC1 values (smaller area and/or lower complexity).

Finally, the *qua1* mutant — known for developing adhesion defects in its epidermis — occupies similar ranges on this chart as the wild type seedlings, although points concentrate towards more elongated and smaller and/or less complex cell shapes.

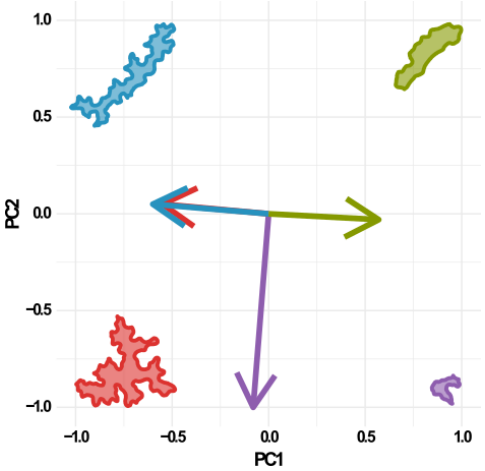
As principal component analysis revealed that the area and the complexity of the cells varies simultaneously, the next step was to examine the actual correlations between these values. Figure 5.4 show area as a function of circularity on a log-log scale, along with the line of best fit. Values of r^2 are presented in the top left corner of every panel and most of them indicate a strong correlation between size and the complexity of shape. Parameters of the line of best fit for each group and for the full wild type sample set are presented in Table 5.3. Recalling the definition of circularity used, we can investigate the mathematical details of this newly found linear relationship on the log-log plot:

$$\log A = a + b \log C = a + b \log \left(\frac{P^2}{4\pi A} \right), \tag{5.1}$$

where A, C and P stand for area, circularity and perimeter respectively and



(a) PCA scores of cells from different lines



(b) PC scores of original descriptors red — area, purple — aspect ratio, blue — circularity, green — solidity. Cell outlines illustrate the variation of cell shapes in the PC1–PC2 morphospace.

Descriptor	PC1	PC2
Area	-0.55	0.05
Aspect ratio	-0.08	-1.00
Circularity	-0.60	0.05
Solidity	0.57	-0.03

Figure 5.3 – PCA on geometric descriptors — PC1 accounts for 65.1%, PC2 covers 24.7% of the variability in the data. PC1 captures information about the absolute area and the complexity of the outline, while PC2 accounts for the elongation of the cells.

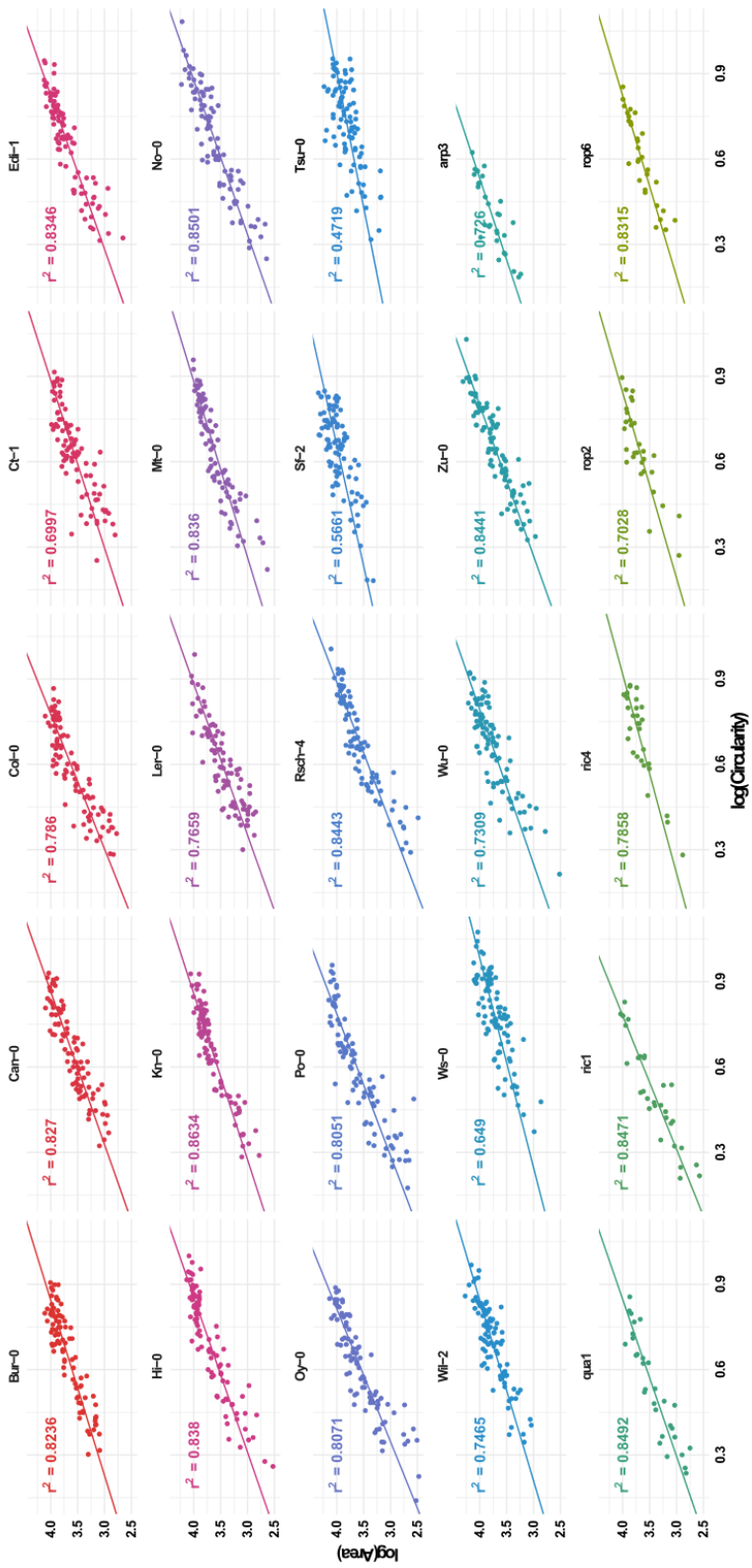


Figure 5.4 – **Correlations between log(area) and log(circularity)** — Line indicates the line of best fit, r^2 indicated on the top left corner of every panel.

line	a	b	r^2	f
Bur-0	2.63	1.63	0.8236	1.62
Can-0	2.39	1.86	0.8270	1.54
Col-0	2.36	2.11	0.7860	1.47
Ct-1	2.50	1.70	0.6997	1.59
Edi-1	2.48	1.85	0.8346	1.54
Hi-0	2.43	1.83	0.8380	1.55
Kn-0	2.53	1.71	0.8634	1.58
Ler-0	2.34	1.88	0.7659	1.53
Mt-0	2.57	1.62	0.8360	1.62
No-0	2.38	1.84	0.8501	1.54
Oy-0	2.26	2.14	0.8071	1.47
Po-0	2.42	2.01	0.8051	1.50
Rsch-4	2.21	2.01	0.8443	1.50
Sf-2	3.22	1.18	0.5661	1.85
Tsu-0	3.05	1.06	0.4719	1.95
Wil-2	2.67	1.57	0.7465	1.63
Ws-0	2.67	1.36	0.6490	1.74
Wu-0	2.54	1.84	0.7309	1.54
Zu-0	2.49	1.89	0.8441	1.53
arp3	3.07	1.75	0.7260	1.57
qua1	2.46	1.82	0.8492	1.55
ric1	2.34	2.13	0.8471	1.47
ric4	2.69	1.43	0.7858	1.70
rop2	2.71	1.54	0.7028	1.65
rop6	2.71	1.58	0.8315	1.63

Table 5.3 – **Linear regression parameters for** $\log A = a + b \log C$
$f = (1 + b)/b$

a and b are the intercept and the slope of the fit line. After exponentiation, this gives

$$A = aC^b = a\left(\frac{P^2}{4\pi A}\right)^b. \tag{5.2}$$

Further rearranging, the proportionality between the area and the perimeter along the fit line is

$$A^{1+b} \propto P^{2b}. \tag{5.3}$$

The area A scales quadratically with length. Using this we can determine how the perimeter P scales as the length L increases:

$$P \propto L^{\frac{1+b}{b}} \tag{5.4}$$

The exponent $(1+b)/b$ is always greater than 1 for positive correlations (b was the slope of the line of best fit). Thus, according to this expression, in order to bring one cell shape into another along the line, the perimeter has to increase more than it would with a true scaling. This indicates that bigger cells in the dataset have proportionally more complex shapes. How the complexity of the outline increases with the area is determined by the exponent.

This scaling behaviour is reminiscent of fractals. The exponent from Equation 5.4, similar to fractal dimension, can be used to quantify the variation of the outline complexity with the area and is presented in Table 5.3 along with the fit parameters.

This property indicates that the growing cell gains area by extending lobes (and in later steps, lobes branching off of these lobes). A pure fractal-like growth pattern would involve initiating new lobes along the perimeter at characteristic time intervals and at a characteristic distance from each other.

Figure 5.5 pictures the first three iterations of the Koch snowflake. At every step, a "lobe" grows in the middle of every edge. The edges of the lobe are a third of the edge length. The last panel shows an *Arabidopsis* cell. Although the cell has a more complex branching structure suggesting that the lobes themselves also grow before new lobes are initiated, there is a certain similarity between the shapes.

This mode of growth is different from what we observed in maize, where the cells increase their area primarily by elongating along the length of the

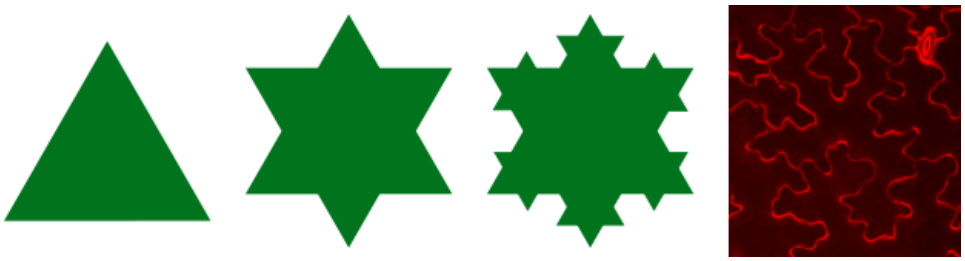


Figure 5.5 – **Fractals and cell shapes**
First three iterations of Koch's snowflake and an *Arabidopsis* cell

leaf and lobe growth does not have a significant contribution to the increase in area.

5.7 Discussion

In this experiment the original intention was to quantify the cell shapes of the nineteen *Arabidopsis* ecotypes selected to be the parents of the MAGIC lines and contrast the differences with the variations in the mutant lines showing cell shape defects. However, the mutant lines were not all so different. The *arp3* line had the least normal cell shapes: different in all four descriptors. Others showing significant differences in most shape features were *ric1* and *qua1*. The difference in circularity between *arp3* and the wild type average is 2.25. Using this to normalise the circularities of the MAGIC parent lines, the median circularity of cell shapes varies in the range of -40%–$+60\%$, with Ler-0 and Po-0 being on the least undulating end and Ws-0 showing the most complex cell shapes. This result, while interesting, has to be treated with care. The variation within a sample is also quite high. The interquartile range of the different samples varies from 90% to 150%. This means that the variability within a sample is at least as high as the variation between samples.

Correlations with geographic locations were also explored but due to the uncertainty in the cell shape data and the fact that raw geographic coordinates are not the most optimal way to represent the climate, no significant correlations were revealed. This does not exclude the possibility that there are correlations with climate factors, however given how the differences within and amongst samples are comparable, results may not have much significance.

Principal component analysis, however, revealed an interesting fact: that contrary to vascular plants in general, the pavement cells of *Arabidopsis* the increase in area (= the growth) is not related to elongation but to reaching a more complex shape. Elongation is only of secondary importance for distinguishing between these cells indicating a lack of global growth anisotropy. Most interestingly, area and circularity seem to be strongly related, indicating a fractalesque growth, which is in line with the observation that this stage of development is when heterotropic growth happens. The similarity between some biological structures and fractal has been observed before and in some areas, fractal dimension is used to quantify shapes [147]. Regarding shapes in plants, McLellan *et al.* mention fractal dimension as a possible parameter to use for their study of *Acer* leaves but they do not use this descriptor in the end.

5.8 Conclusions

In this chapter, the geometric descriptor method was used to quantify the cell shapes of the 19 MAGIC parent lines and 6 mutant lines known for showing cell shape defect to a varying extent. It was found that the geometric descriptors work in this case too, but with *Arabidopsis* the important quantities were area and circularity.

Differences between the MAGIC parents lines were, with a few exceptions, found not significant: the spread within the datasets is comparable to the differences of the median values. Some of the mutant lines were found to be significantly different, confirming that the method works.

Finally, an interesting relationship was confirmed between area and the perimeter of *Arabidopsis* cells within the same leaf, showing fractal-like scaling and indicating that cells with larger areas do not expand as much isotropically as they grow heterotropically. This property was already known from cell tracking experiments and by theoretical speculations, but deriving it from the results was a pleasant surprise.

Chapter 6

Discussion

6.1 Sampling

As presented in Chapters 3–4, epidermal cell shapes change with the age of the plant, depending on the position, the side and the number of the leaf, and of course there is a great variance between species. Although not studied over the course of this project but it is also known that environmental factors can affect the shape.

With so many influencing factors, a well-chosen sampling strategy is essential to optimise the effort required and the quality of the results produced. The maize dataset most certainly suffers from an insufficient sampling routine. In retrospect, it would have been a better choice to study the cell shapes on the first $\approx 10\,\mathrm{cm}$ of the leaves in a consecutive manner, rather than selecting a few evenly spaced segments along the full length, as much of the shape change occurs in this region. Measuring the distance from the ligule is adequate but a better accuracy is necessary when cutting samples.

Regarding the vascular plants dataset, apart from monocots and eudicots, the remaining groups contain very few species (although the proportions may represent the total number of species pertaining to each group better). The current fern samples seem to exhibit a very wide range of shapes, with very different mean values. It could be interesting to explore whether there is a real variation or if the four species selected contain some outliers.

With *Arabidopsis*, although the data hints at an interesting growth phenomenon by relating shape to the absolute area, this is merely speculation as

all the data was collected at a single time point (7 days after germination). This result could be enhanced by also studying the shape changes over a period of time. In *Arabidopsis*, thanks to the very flat cotyledons, it is even possible to track the growth of the same cell.

6.2 Sample preparation and imaging

Sample preparation is easy to standardise when the subject of the experiment is a single species. Although the leaf clearing technique used on samples of the vascular plant dataset is very widely applicable, the effect of the toluidine blue staining varies quite significantly between species. Most often the staining worked well, but the background colour and intensity varies noticeably.

Throughout this project different imaging techniques were tried in the hope of obtaining images suitable for automatic segmentation. For this, it is helpful to capture images with similar contrast and overall colour levels. In this respect, confocal images proved to work best.

However, confocal microscopy also presents a set of issues. The efficiency of staining very much depends on the species: *Arabidopsis* cotyledons are fully stained after little more than a minute in the propidium iodide solution, while with maize leaves only the edges of the segment stain, even after 5 minutes of vacuum infiltration. In order to stain larger areas of the epidermis (of order 10 cells), one can try gently scratching the bottom surface of the leaf, but this often leads to uneven staining. Moreover, as discussed before, even with a well-stained sample, the lobes of maize cells are difficult to image in a confocal microscope (or by any other method that can image inside the tissue).

The staining issue can be circumvented by using fluorescent lines. Such lines are only available for model species like *Arabidopsis* or maize. These lines are created to mark particular features — for example, the cell wall of the cell — by a fluorescent protein. This would be extremely useful for obtaining better images for the maize dataset. Confocal microscopy is also capable to capture 3D cell shapes.

6.3 3D shapes

Epidermal cell shapes are usually studied in only two dimensions. Although adding an extra dimension would clearly render the problem even more complex, it would also be interesting to see which species develop lobes along the full length of the anticlinal wall, like *Arabidopsis*, and which species produce superficial lobes, like maize. While collecting images for the vascular dataset, superficial lobing similar to maize was suspected in other members of the Poaceae family. The optical microscope used here also allowed imaging at different z depths, like confocal microscopy, but because of the strong staining of the cell walls it was not possible to obtain images at a depth where the lobes disappear.

6.4 Segmentation

In these experiments, segmentation was the real limiting factor on the quantity of data that could be analysed. In order to handle larger datasets, segmentation must be automatised. This is especially true for experiments that involve tracking the shape changes of selected cells (feasible in *Arabidopsis*) or in experiments studying 3D cell shapes. Automatic segmentation tools are more widely available for confocal images. A popular choice is MorphoGraphX [148], which can also perform 3D segmentation on stacks of images.

6.5 Quantitative methods

Elliptic Fourier descriptors can provide very good descriptions of the shapes at a sufficiently high harmonic number and carry the advantage that cell shapes can be reconstructed from the descriptors. Elliptic Fourier descriptors were trialled on the maize dataset and also used on the vascular plants. In both cases, the first few principal components failed to capture the undulations in the shape of the outline.

Instead, the first principal component typically picked out elongation. In both cases, two of the next three components represented rotational symmetry and inversion (PC3 and PC4 for maize and PC2 and PC3 in the vascular plant dataset). The complex undulations (described by higher harmonics)

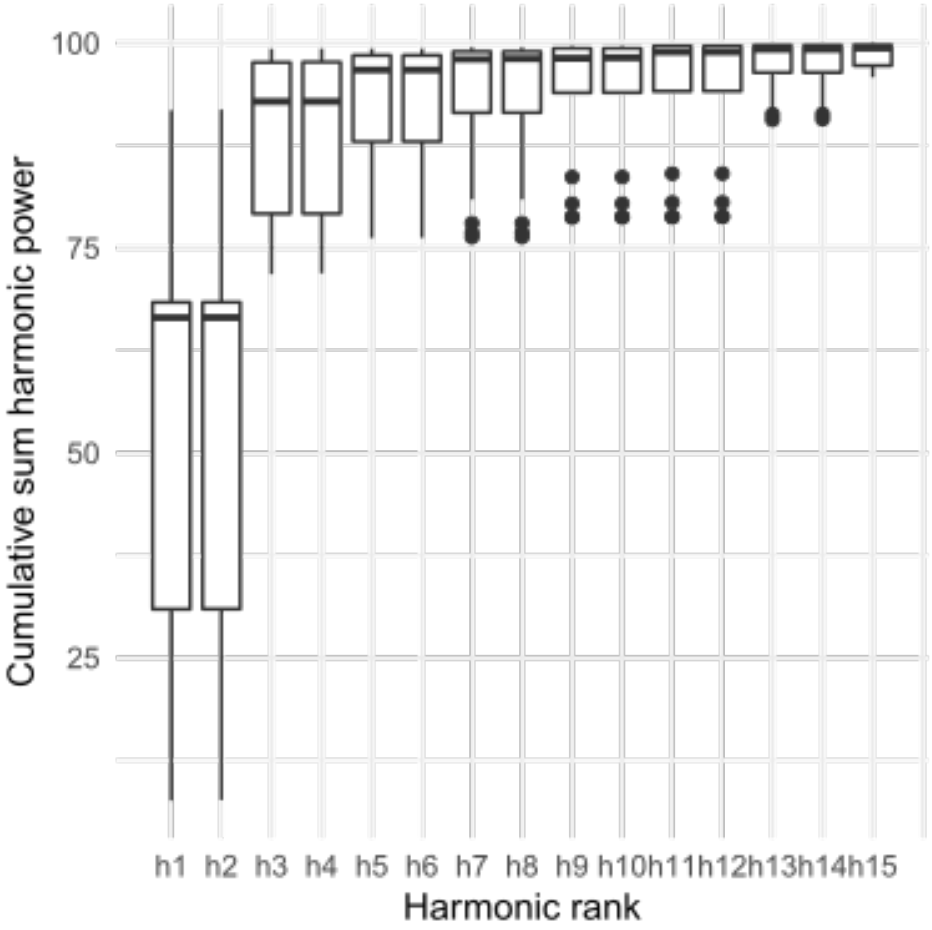
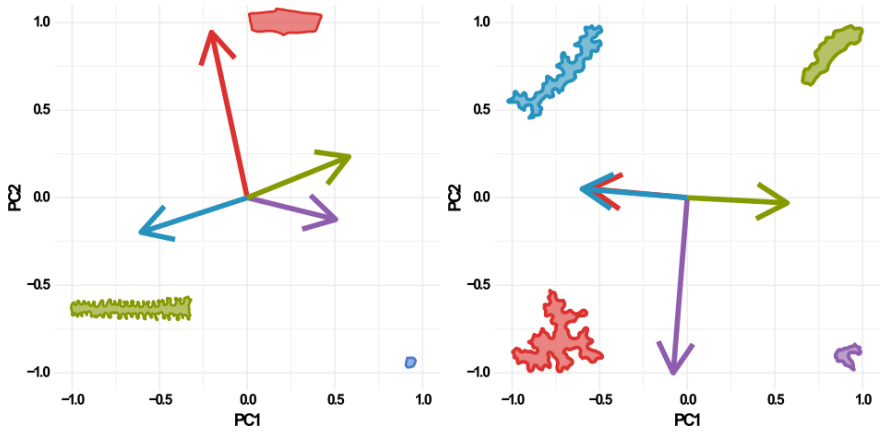


Figure 6.1 – Cumulative contributions of the first 15 harmonics for the set of example cell shapes used in the analysis in Chapter 2.

were never present in the first four principal components.

The disappearance of outline complexity can be explained by two factors: firstly, in a real dataset, the lobes on the cells are more randomly located and lobe number vary as well. This means that the outline information is shared between several higher harmonic components instead of a particular harmonic as in case of the fairly symmetric example shapes used in Chapter 2. Secondly, the contribution of higher harmonics typically decreases with harmonic number — see Figure 6.1 showing the cumulative contributions of the first 15 harmonics for the set of example shapes. The reason for this is that the contribution is calculated as the contribution to the area captured by the reconstructed shape, and not the perimeter.



(a) Maize **(b)** *Arabidopsis*

Figure 6.2 – Comparing the contribution of geometric descriptors for maize and *Arabidopsis*
Area (red), aspect ratio (purple), circularity (blue) and solidity (green) shown in the PC1–PC2 morphoplane for maize and *Arabidopsis*. Example cells illustrate how variations affect the cell shape.

Fourier analysis can also be modified to provide a more meaningful basis set [149], while retaining the important feature of Fourier analysis that it allows a perfect reconstruction of the cell shapes.

Geometric descriptors have, in all three cases, provided a good description of the datasets, although the relative importance of area, aspect ratio, circularity and solidity varied. Solidity was found to be a better descriptor for quantifying outline complexity in maize and also in the vascular plant dataset alongside aspect ratio used to quantify elongation. The particular shape of *Arabidopsis* cells — low variation in elongation but high variation in outline complexity — meant that circularity was a better descriptor of outline complexity as it could distinguish simple lobes and branching lobes (similar to rows 3 and 4 of the example shapes shown on Figure 2.3).

Although geometric descriptors only included 4 variables, principal component analysis was still used to define a two-dimensional morphospace that captures the largest possible variation in the dataset. Using the first two principal components allowed studying the variations in the dataset and also provided information about how dependent or independent the four descriptors were when projected onto this plane.

The directions along which area, aspect ratio, circularity and solidity vary in the PC1–PC2 morphospace are shown on Figure 6.2 for the maize and

the *Arabidopsis* dataset. Examining the position of the shape descriptors in relation to the area is particularly interesting: equating the direction in which the area increases with developmental time can tell us about the main mode of cell expansion.

In maize, the directions along which solidity and circularity vary are nearly perpendicular to the area increase. This means that variations in the outline do not have a significant contribution to increasing the area. Aspect ratio, however, decreases as the area increases. This means that elongation (decreasing aspect ratio) is the main contributor to area increase. Aspect ratio is also related to the outline descriptors: as the cell elongates and the aspect ratio decreases, the outline complexity increases — solidity decreases and circularity increases.

In *Arabidopsis*, an increase in area is related to an increase in outline complexity, suggesting that lobe growth is the primary way through which the cell expands. Aspect ratio varies independently of area and the shape descriptors, indicating that the global anisotropy of cell expansion is insignificant.

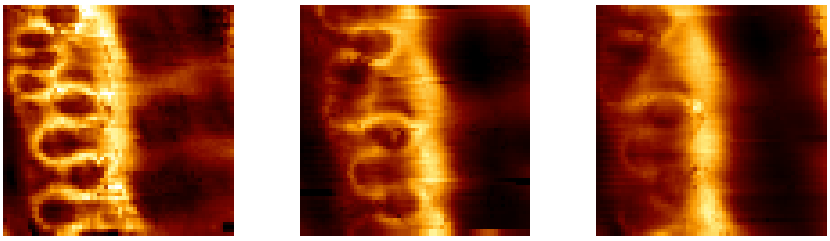
Although it is interesting to derive and compare modes of cell expansion from geometric descriptors, only the maize dataset contains cell shapes from plants at a different age. In *Arabidopsis*, I merely used the assumption that bigger cells are likely to be further ahead in terms of cell expansion. Therefore this finding would have to be verified on a dataset containing cells from plants at different ages, preferably on a dataset tracking the expansion of individual cells over the course of several days.

6.6 Shape and strength

This project originally also intended to explore some factors that may contribute to the particular cell shapes observed. One such factor is the elasticity of the anticlinal cell wall along the outline of the cell.

Maize leaf segments were studied in an Atomic Force Microscope (AFM) in force mapping mode. The force mapping mode allows measuring the (visco)elastic properties from the response to a vertical indentation by a force on the nN scale. Leaf segments were kept in a highly concentrated mannitol solution before and also during the measurements to remove the turgor pressure. The force-displacement curve was recorded at several points

and from this data the elastic modulus was calculated for the target area. Figure 6.3 shows the magnitude of the elastic modulus for the same area taken at an indentation force of 2000, 5000 and 7000 nN respectively.



(a) 2000 nN **(b)** 5000 nN **(c)** 7000 nN

Figure 6.3 – Maps of the elastic modulus using different indentation forces — greater force corresponds to 'seeing' deeper. These images prove that the lobes only exists at the surface of the epidermis. This also means that the so calculated elastic moduli at different points cannot be compared as the curvature of the cell wall varies significantly with the location.

Greater force corresponds to a greater indentation depth. In the case of maize, this experiment revealed that the lobes are only present on the surface. This also meant that the anticlinal cell wall is not vertical and therefore the models used to estimate the elastic modulus were not applicable. Moreover, at 2000 nN the data was very noisy. Increasing the force improved the reliability of the data but the experiment was no longer capturing the elasticity of the cell wall on the edges of the lobes, therefore this attempt was abandoned.

6.7 Shape reproducibility

Another interesting experiment was planned originally to examine the effect of overall mechanical conditions on the shape growth of pavement cells. The idea was to first obtain protoplasts of epidermal pavement cells by digesting away the cell wall and then order these round cells into a single cell layer either by floating and packing on liquids (Langmuir-Blodgett method, [150]) or by 3D-printing them onto a stretchable substrate [151] and let them regrow the cell wall under different mechanical conditions. This would have allowed comparing the obtained cell outlines to the ones in the original tissue and maybe to quantify mechanical conditions during growth to potentially explain why some species develop undulations while others do not.

The experiment was originally designed with maize protoplasts in mind. Maize protoplasts were found to be quite vulnerable to work with and later I decided to experiment with *Arabidopsis* instead, as working with *Arabidopsis* protoplasts is much better documented. Even so, it turned out to be really hard to keep protoplasts alive for long enough. As well as using the Braybrook lab protocols, I also borrowed media recipes from Pau Formosa-Jordan. Although I have dedicated the better part of summer 2015 to perfecting the art of isolating and culturing protoplasts, I have not managed to keep them alive for longer than 2–4 days under sterile conditions and so I gave up on this idea.

6.8 Underlying mechanical properties

Since this thesis was first submitted, a very interesting work has been published in which first the mechanical stress endured by differently shaped cells was modelled [152]. Results showed that expanding in one direction (elongation) did not increase the mechanical stress in the cell wall. When the cell was expanding in two or three dimensions, the mechanical stress experienced by the cell was has, however, drastically increased.

Sapala *et al.* hypothesised that cells may be growing heterotropically (through lobes extending in multiple directions) in situations where the tissue as a whole is expanding isotropically, and that lobes are formed in order to reduce the mechanical stress experienced by the cell wall. This idea was tested by modelling the expansion of a set of cells whose shape was extracted from confocal images of *Arabidopsis* cotyledons under a range of conditions: overall isotropic and anisotropic tissue expansion, in a non-uniformly expanding tissue and also varied parameters such as the bending stiffness of the cell wall, the stiffness of the cellulose fibres for example.

The cell outlines obtained in these simulations bear a shockingly familiar resemblance to what I observed during my sampling of the eudicots. The authors also examined a number of species with puzzle-like or non-lobed cells and found that "lobeyness" (defined as the ratio of the perimeter to the perimeter of the convex hull) correlates with the area of the cells across species. This is very similar to the results I obtained in the *Arabidopsis* dataset but differs from the overall relationship I obtained for vascular plants, where area showed a stronger overall correlation with aspect ratio

due to the presence of elongated cells.

Although this model also successfully explains why cells might don a simple elongated shape, it does not explain the cell shapes observed in maize and other members of the Poales, where the cells are heavily elongated and lobes appear only towards the surface and not at the bottom part of the epidermal cell layer.

Chapter 7

Conclusions and outlook

In this project, three extensive datasets of cell shapes were collected and quantitatively analysed. The maize dataset contains 1620 cells, similar in magnitude to the *Arabidopsis* dataset with 1890 individual cells. The vascular plant dataset is by far the largest, with 10260 cells collected from 230 species.

Methods for quantification were tested and it was found that in all of these cases, elliptic Fourier descriptors are inferior to geometric descriptors (area, aspect ratio, circularity and solidity) in resolving these points.

Data from vascular plants was analysed in a phylogenetic context and revealed differences in the phylogenetic signal between different clades. Regarding the undulation of the cell outlines, it was found that not only the averages can be different between species, but that the range of solidities within a species varies too in eudicot families. The differences between ad- and abaxial sides were also more similar within families.

Elongation and the undulation depth of the outline were important characteristics in maize and overall with the vascular plants, these were measured by aspect ratio and solidity respectively. With *Arabidopsis*, aspect ratio did not seem to matter anymore and here circularity was a stronger descriptor than solidity. Circularity showed an interesting correlation with area, providing mathematical evidence for the fractalesque growth of *Arabidopsis* cells. In maize, the relationship between elongation and undulation depth was found to be different for the juvenile and adult phases of development, numerically confirming a long-standing observation.

Both of these observations correspond to modes of growth that minimise

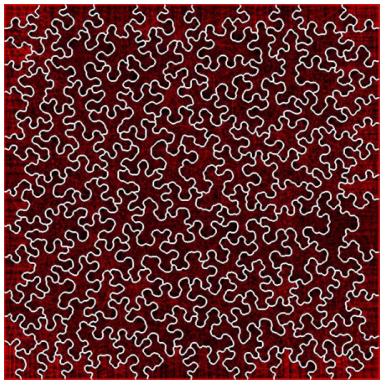
the mechanical stress on the cell wall [152], but under different overall tissue expansion conditions: maize leaves are highly anisotropic and so the cells grow preferentially by elongating parallel with the length of the leaf. *Arabidopsis* leaves expand isotropically, but in order to keep the mechanical stress on the cell wall at a minimum (corresponding to keeping the minimum cross-section somewhat constant), the cells expand lobes. As the lobed cell would expand further isotropically, new lobes are initiated instead in order to preserve the minimal stress condition. This is what gives rise to the fractal-type correlation between area and shape complexity: attempting to expand along one dimension only (in a local sense) while filling a two-dimensional area. Fractals can be thought of as a way to fill a higher-dimensional space by a lower dimensional curve.

The lobing in maize remains unexplained: as these lobes appear only close to the surface, only 3D modelling could elucidate why it is favourable to expand this way. Moreover, not only cells but also leaves exist in three dimensions: maize leaves are, for example, wound around the stem at the bottom of the plant and start bending under gravity when they reach a certain length. Undoubtedly, the bending of the leaf also modifies the mechanical stress experienced by the leaves. Could this be responsible for the superficial lobes observed in the Poales?

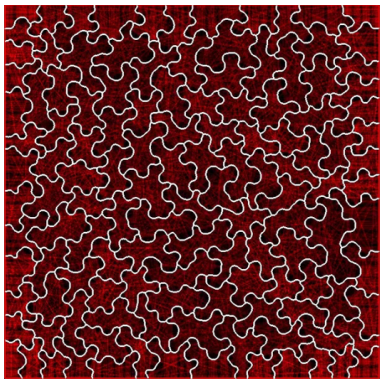
It would be interesting to see the results from a model considering the mechanical context of an expanding leaf and its epidermal cells in 3D, and whether it explains the presence of lobes on a maize cell or if the response lies elsewhere. 3D modelling would also be necessary to explain the often-present difference between adaxial and abaxial sides: according the observations presented in Chapter 4, cells on the abaxial side are often more lobed than on the adaxial side, however the area also tends to be greater on the adaxial side than on the abaxial side, although the distribution is less skewed. Although the mechanical properties on a cellular level could be responsible for these differences, it would be worth considering the effect of the 3D shape of the leaf on the cells in the adaxial and abaxial sides.

As the model presented in [152] is based purely on mechanical properties and yet the results resemble the fine variation observed in eudicot species, it would be even more essential to show experimentally that it really is the mechanical context that matters, and not the genetic background (provided, of course, that the necessary RIC and ROP proteins enabling interdigitation

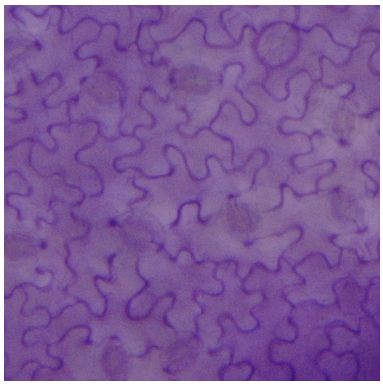
are present). This could possibly involve a strategy similar to the one described in Section 6.7 — removing protoplasts from the epidermis and forcing them to grow into shapes different from their natural shapes. Alternatively or additionally, measurements of mechanical properties of the cell wall and the resulting cell shapes could be compared to the outlines derived in the model presented in [152] to explore the extent to which they describe the subtle variances in patterning between species. Figure 7.1 shows two images from [152] of cell wall outlines obtained at different bending stiffness values used in the simulation against images of the epidermal cells in *Pastinaca sativa* and *Galium odoratum*, presenting very similar cell patterning.



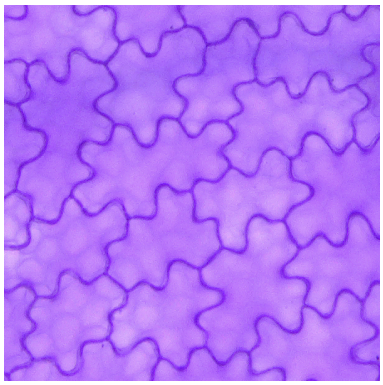
(a) Simulated epidermal cell patterning at lower bending stiffness



(b) Simulated epidermal cell patterning at higher bending stiffness



(c) *Pastinaca sativa* abaxial epidermis, taken at 700× magnification



(d) *Galium odoratum* adaxial epidermis, taken at 400× magnification

Figure 7.1 – Theory and experiment in apparent accord: varying the bending stiffness of the cell wall in the simulation yields cell patterning resembling different species. Simulated images taken from [152].

Chapter 8

Acknowledgements

This thesis could not have been written without the support I received from many people over the course of the past three years.

I thank Siobhan Braybrook for opening my eyes to the beautiful diversity of plant cell shapes and the Braybrook group for all their help during my time with them.

I am grateful for having access to the excellent imaging facilities of the Sainsbury Laboratory, the collection of the Cambridge University Botanic Garden and the various libraries across Cambridge possessing an extreme wealth of botanical literature from past centuries.

I benefited greatly from being part of the NanoDTC, first and foremost by allowing me to meet so many inspirational scientists across all areas.

I am eternally indebted to Trinity College, for being my true home for so many years and for providing so many opportunities and friends. I met many wonderful minds within the BA Society and I am especially grateful to members of the BA Committees of 2015-2016 and 2016-2017 — we have achieved so much together.

I would like to especially thank the following people for their help with this project:

Sally Petit and Alex Summers from CUBG, for theis help with collecting samples from the Botanic Garden;

Madelaine Bartlett and her students, Grace Pisano, Rebecca Goldberg and Jeffrey Heithmar (UMass Amherst), for their part in our collaboration to study the cell shapes in vascular plants;

Firas Bou Daher (SLCU), for his help in planning my *Arabidopsis* experiment, training me on the confocal microscope and providing the seeds of the mutant lines;

Katie Abley, for providing seeds of the MAGIC parent lines;

Marco Aita and Louis Kaplan (SLCU), for in-depth mathematical discussions and general merryment;

Alexis Daboville, for his help with programming concepts and thesis aesthetics; and

Ádám Tamás Nagy, for latin language help.

I also thank my examiners, Veronica Grieneisen and Henrik Jönsson, for their helpful comments and suggestions.

Bibliography

[1] Geitmann, A. and Ortega, J. K. *Trends in Plant Science* **14**(9), 467 – 478 (2009).

[2] Hooke, R. *Micrographia, or, Some physiological descriptions of minute bodies made by magnifying glasses with observations and inquiries thereupon.* Printed by J. Martyn and J. Allestry, (1665).

[3] Plateau, J. A. F. *Statique experimentale et theorique des liquides soumis aux seules forces moleculaires.* Gauthier-Villars, Paris, (1873).

[4] Taylor, J. E. *Annals of Mathematics* **103**(3), 489–539 (1976).

[5] Thomson, S. W. *Philosophical Magazine* **24**(151), 503–514 (1887).

[6] Lewis, F. T. *American Journal of Botany* **30**(1), 74–81 (1943).

[7] Marvin, J. W. *American Journal of Botany* **26**(5), 280–288 (1939).

[8] Matzke, E. B. *Proceedings of the National Academy of Sciences of the United States of America* **31**(9), 281–289 (1945).

[9] Treviranus, L. C. *Physiologie der Gewächse.* A. Marcus, (1838).

[10] Linsbauer, K. *Handbuch der Pflanzenanatomie* **4**(1), 134–171 (1930).

[11] Areschoug, F. W. C. *Über die physiologischen Leistungen und die Entwickelung des Grundgewebes des Blattes.* E. Malmströms Buchdruckerei, (1897).

[12] Anheisser, R. *Flora* **87**, 64 (1900).

[13] Watson, R. W. *New Phytologist* **41**(4), 223–229 (1942).

[14] Brenner, W. *Flora* **87**, 387 (1900).

[15] Ambronn, H. *Jb. wiss. Bot.(14)* , 82–110 (1884).

[16] Haberlandt, G. *Jb. wiss. Bot.(13)* , 74–188 (1882).

[17] Pfeffer, W. *The Physiology Of Plants. Vol. 2.* The Clarendon Press; Oxford, (1903).

[18] Lee, B. and Priestley, J. H. *Annals of Botany* **38**(151), 525–545 (1924).

[19] Avery, G. S. *American Journal of Botany* **20**(9), 565–592 (1933).

[20] Dodoens, R. and de L'Écluse, C. *Histoire des plantes...* Biblioteca digital Dioscórides. de l'Imprimerie de Iean Loë, (1557).

[21] Askenasy, E. *Bot. Ztg* **28**, 193–209 (1870).

[22] Mommer, L. and Visser, E. J. W. *Annals of Botany* **96**(4), 581–589 (2005).

[23] Korn, R. W. *New Phytologist* **73**(5), 927–935 (1974).

[24] Jenks, M. A., Rashotte, A. M., Tuttle, H. A., and Feldmann, K. A. *Plant Physiology* **110**(2), 377–385 (1996).

[25] Sylvester, A. W., Parker-Clark, V., and Murray, G. A. *American Journal of Botany* **88**(12), 2157–2167 (2001).

[26] Poethig, R. S. *Science* **301**(5631), 334–336 (2003).

[27] Bongard-Pierce, D. K., Evans, M. M. S., and Poethig, R. S. *International Journal of Plant Sciences* **157**(4), 331–340 (1996).

[28] Hales, S. *Vegetable Staticks: Or, an Account of Some Statical Experiments on the Sap in Vegetables: Being an Essay Towards a Natural History of Vegetation. Also, a Specimen of an Attempt to Analyse the Air, by a Great Variety of Chymio-statical Experiments; which Were Read at Several Meetings Before the Royal Society.* W. and J. Innys and T. Woodward, (1727).

[29] Sachs, J. *Lehrbuch der Botanik nach dem gegenwärtigen Stand der Wissenschaft.* W. Engelmann, (1874).

[30] Pfeffer, W. *Pflanzenphysiologie: ein Handbuch der Lehre vom Stoffwechsels und Kraftwechsels in der Pflanze.* W. Engelmann, (1904).

[31] Peters, W. S. and Tomos, A. D. *Annals of Botany* **77**(6), 657–665 (1996).

[32] Kutschera, U. and Niklas, K. *Journal of Plant Physiology* **164**(11), 1395 – 1409 (2007).

[33] Kutschera, U. *Annals of Botany* **101**(5), 615–621 (2008).

[34] Roelofsen, P. volume 2 of *Advances in Botanical Research*, 69 – 149. Academic Press (1966).

[35] Takeda, K. and Shibaoka, H. *Planta* **151**(4), 393–398 (1981).

[36] Paolillo, D. J. *New Phytologist* **145**(3), 449455 (2000).

[37] Baskin, T. I. and Jensen, O. E. *Journal of Experimental Botany* **64**(15), 4697–4707 (2013).

[38] Wasteneys, G. O. and Galway, M. E. *Annual Review of Plant Biology* **54**(1), 691–722 (2003).

[39] Fu, Y., Gu, Y., Zheng, Z., Wasteneys, G., and Yang, Z. *Cell* **120**(5), 687–700 (2005).

[40] Zhang, C., Halsey, L. E., and Szymanski, D. B. *BMC Plant Biology* **11**(1), 27 (2011).

[41] Paredez, A. R., Somerville, C. R., and Ehrhardt, D. W. *Science* **312**(5779), 1491–1495 (2006).

[42] Frank, M. J. and Smith, L. G. *Current biology* **12**(10), 849–853 (2002).

[43] Fu, Y., Li, H., and Yang, Z. *The Plant Cell* **14**(4), 777–794 (2002).

[44] Ivakov, A. and Persson, S. *Frontiers in Plant Science* **4**, 439 (2013).

[45] Machesky, L. M. and Gould, K. L. *Current Opinion in Cell Biology* **11**(1), 117 – 121 (1999).

[46] H M. *Cell* **76**(3), 555 – 566 (1994).

[47] Li, S., Blanchoin, L., Yang, Z., and Lord, E. M. *Plant Physiology* **132**(4), 2034–2044 (2003).

[48] Mathur, J., Mathur, N., Kirik, V., Kernebeck, B., Srinivas, B. P., and Hülskamp, M. *Development* **130**(14), 3137–3146 (2003).

[49] Mathur, J., Mathur, N., Kernebeck, B., and Hülskamp, M. *The Plant Cell* **15**(7), 1632–1645 (2003).

[50] El-Assal, S. E.-D., Le, J., Basu, D., Mallery, E. L., and Szymanski, D. B. *Current Biology* **14**(15), 1405 – 1409 (2004).

[51] Frank, M., Egile, C., Dyachok, J., Djakovic, S., Nolasco, M., Li, R., and Smith, L. G. *Proceedings of the National Academy of Sciences of the United States of America* **101**(46), 16379–16384 (2004).

[52] Gautreau, A., Hsin-yi, H. H., Li, J., Steen, H., Gygi, S. P., and Kirschner, M. W. *Proceedings of the National Academy of Sciences of the United States of America* **101**(13), 4379–4383 (2004).

[53] Basu, D., El-Assal, S. E.-D., Le, J., Mallery, E. L., and Szymanski, D. B. *Development* **131**(17), 4345–4355 (2004).

[54] Basu, D., Le, J., El-Essal, S. E.-D., Huang, S., Zhang, C., Mallery, E. L., Koliantz, G., Staiger, C. J., and Szymanski, D. B. *The Plant Cell* **17**(2), 502–524 (2005).

[55] Brembu, T., Winge, P., Seem, M., and Bones, A. M. *The Plant Cell* **16**(9), 2335–2349 (2004).

[56] Deeks, M. J., Kaloriti, D., Davies, B., Malh, R., and Hussey, P. J. *Current Biology* **14**(15), 1410 – 1414 (2004).

[57] Zhang, X., Dyachok, J., Krishnakumar, S., Smith, L. G., and Oppenheimer, D. G. *The Plant Cell* **17**(8), 2314–2326 (2005).

[58] Djakovic, S., Dyachok, J., Burke, M., Frank, M. J., and Smith, L. G. *Development* **133**(6), 1091–1100 (2006).

[59] Le, J., Mallery, E. L., Zhang, C., Brankle, S., and Szymanski, D. B. *Current Biology* **16**(9), 895 – 901 (2006).

[60] Bichet, A., Desnos, T., Turner, S., Grandjean, O., and Höfte, H. *The Plant Journal* **25**(2), 137–148 (2001).

[61] Whittington, A. T., Vugrek, O., Wei, K. J., Hasenbein, N. G., et al. *Nature* **411**(6837), 610 (2001).

[62] Kotzer, A. and Wasteneys, G. *Botany* **84**(4), 594–603 (2006).

[63] Kirik, V., Herrmann, U., Parupalli, C., Sedbrook, J. C., Ehrhardt, D. W., and Hülskamp, M. *Journal of Cell Science* **120**(24), 4416–4425 (2007).

[64] Wang, X., Zhu, L., Liu, B., Wang, C., Jin, L., Zhao, Q., and Yuan, M. *The Plant Cell* **19**(3), 877–889 (2007).

[65] Wasteneys, G. O. and Ambrose, J. C. *Trends in Cell Biology* **19**(2), 62 – 71 (2009).

[66] Lin, D., Cao, L., Zhou, Z., Zhu, L., Ehrhardt, D., Yang, Z., and Fu, Y. *Current Biology* **23**(4), 290–297 (2013).

[67] Sampathkumar, A., Krupinski, P., Wightman, R., Milani, P., Berquand, A., Boudaoud, A., Hamant, O., Jönsson, H., and Meyerowitz, E. M. *Elife* **3**, e01967 (2014).

[68] Fu, Y., Xu, T., Zhu, L., Wen, M., and Yang, Z. *Current Biology* **19**(21), 1827–1832 (2009).

[69] Burk, D. H., Liu, B., Zhong, R., Morrison, W. H., and Ye, Z.-H. *The Plant Cell* **13**(4), 807–827 (2001).

[70] Jacques, E., Verbelen, J.-P., and Vissenberg, K. *BMC Plant Biology* **13**(1), 163 (2013).

[71] Xu, T., Wen, M., Nagawa, S., Fu, Y., Chen, J.-G., Wu, M.-J., Perrot-Rechenmann, C., Friml, J., Jones, A. M., and Yang, Z. *Cell* **143**(1), 99 – 110 (2010).

[72] Nagawa, S., Xu, T., Lin, D., Dhonukshe, P., Zhang, X., Friml, J., Scheres, B., Fu, Y., and Yang, Z. *PLOS Biology* **10**(4) (2012).

[73] Mettenius, G. *Über die Hymenophyllaceae*. S. Hirzel, (1865).

[74] Panteris, E., Apostolakos, P., and Galatis, B. *New Phytologist* **127**(4), 771–780 (1994).

[75] Armour, W. J., Barton, D. A., Law, A. M., and Overall, R. L. *The Plant Cell* **27**(9), 2484–2500 (2015).

[76] Kendall, D. G. *Bulletin of the London Mathematical Society* **16**(2), 81–121 (1984).

[77] Magono, C. and Lee, C. W. *Journal of the Faculty of Sciences, Hokkaido University* **2**(4), 321–335 (1966).

[78] Wang, P. and Denzer, S. *Journal of the Atmospheric Sciences* **40**(4), 1024–1028 (1983).

[79] Wang, P.-K. *Journal of the Atmospheric Sciences* **39**(11), 2615–2622 (1982).

[80] Zelditch, M., Swiderski, D., and Sheets, H. *Geometric Morphometrics for Biologists: A Primer*. Academic Press. Elsevier Academic Press, (2012).

[81] Bookstein, F. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**(6), 567–585 (1989).

[82] Thompson, D. *On Growth and Form*. Dover Books on Biology Series. Dover, (1942).

[83] Arthur, W. *Nature Reviews Genetics* **7**(5), 401–406 (2006).

[84] Jensen, R. J., Ciofani, K. M., and Miramontes, L. C. *Taxon* **51**(3), 475–492 (2002).

[85] Ray, T. S. *American Journal of Botany* **79**(1), 69–76 (1992).

[86] McLellan, T., Endler, J. A., and Mabee, P. *Systematic Biology* **47**(2), 264–281 (1998).

[87] Peura, M. and Iivarinen, J. In *Proceedings of the third international workshop on visual form*, volume 443, 451, (1997).

[88] Carpenter, A. E., Jones, T. R., Lamprecht, M. R., Clarke, C., Kang, I. H., Friman, O., Guertin, D. A., Chang, J. H., Lindquist, R. A., Moffat, J., et al. *Genome biology* **7**(10), R100 (2006).

[89] Zhang, D. and Lu, G. *Pattern Recognition* **37**(1), 1 – 19 (2004).

[90] Bonhomme, V., Picq, S., Gaucherel, C., and Claude, J. *Journal of Statistical Software* **56**(13) (2014).

[91] K W. M. *Review of Palaeobotany and Palynology* **96**(1), 1 – 30 (1997).

[92] Barrett, A. M. *The Journal of Pathology and Bacteriology* **86**(1), 9–20 (1963).

[93] Young, I. T., Walker, J. E., and Bowie, J. E. *Information and Control* **25**(4), 357 – 370 (1974).

[94] Landau, L., Lifshitz, E., Kosevich, A., and Pitaevskiĭ, L. *Theory of Elasticity*. Course of theoretical physics. Butterworth-Heinemann, (1986).

[95] Hamant, O. and Moulia, B. *New Phytologist* **212**(2), 333–337 (2016).

[96] Hu, M.-K. *IRE Transactions on Information Theory* **8**(2), 179–187 February (1962).

[97] Khotanzad, A. and Hong, Y. H. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**(5), 489–497 (1990).

[98] Zernike, F. *Monthly Notices of the Royal Astronomical Society* **94**, 377–384 (1934).

[99] Pincus, Z. and Theriot, J. A. *Journal of Microscopy* **227**(2), 140–156 (2007).

[100] Fourier, J. *Nouveau Bulletin des Sciences de la Société Philomathique de Paris* **6**, 112–116 (1808).

[101] Kuhl, F. P. and Giardina, C. R. *Computer graphics and image processing* **18**(3), 236–258 (1982).

[102] Zahn, C. T. and Roskies, R. Z. *IEEE Transactions on Computers* **C-21**(3), 269–281 March (1972).

[103] Cosgriff, R. *Ohio State University Research Foundation, Columbus* , Report No. 820–11 (1960).

[104] Pauwels, E. J., de Zeeuw, P. M., and Ranguelova, E. B. *Engineering Applications of Artificial Intelligence* **22**(1), 26 – 31 (2009).

[105] Wu, T.-C., Belteton, S. A., Pack, J., Szymanski, D. B., and Umulis, D. M. *Plant Physiology* **171**(4), 2331–2342 (2016).

[106] Li, M., Frank, M. H., Coneva, V., Mio, W., Topp, C. N., and Chitwood, D. H. *bioRxiv* (2017).

[107] Nelissen, H., Gonzalez, N., and Inzé, D. *Current Opinion in Plant Biology* **33**(Supplement C), 72 – 76 (2016).

[108] Nelissen, H., Rymen, B., Jikumaru, Y., Demuynck, K., Lijsebettens, M. V., Kamiya, Y., Inz, D., and Beemster, G. T. *Current Biology* **22**(13), 1183 – 1187 (2012).

[109] Poethig, R. S. *Science* **250**(4983), 923–930 (1990).

[110] Goebel, K. *Organography of Plants: General organography.* Clarendon, (1900).

[111] Sharman, B. C. *Annals of Botany* **6**(22), 245–282 (1942).

[112] Sylvester, A., Cande, W., and Freeling, M. *Development* **110**(3), 985–1000 (1990).

[113] Muller, B., Reymond, M., and Tardieu, F. *Journal of Experimental Botany* **52**(359), 1259–1268 (2001).

[114] Lawson, E. J. and Poethig, R. S. *Trends in Genetics* **11**(7), 263 – 268 (1995).

[115] Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. *Nature Methods* **9**(7), 671–675 (2012).

[116] Kimball, S. and Mattis, P. *GNU Image Manipulation Program 2.08.10*, (2013).

[117] Pound, M. P., French, A. P., Wells, D. M., Bennett, M. J., and Pridmore, T. P. *The Plant Cell* **24**(4), 1353–1361 (2012).

[118] The MathWorks, Inc., Natick, Massachusetts, United States.

[119] R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, (2018).

[120] Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, (2009).

[121] Freeling, M. *Developmental Biology* **153**(1), 44 – 58 (1992).

[122] Sam, O., Jeréz, E., and Varela, M. *Cultivos tropicales* **17**(2), 32–38 (1996).

[123] Greguss, P. *Acta Biol.(Szeged)* **3**, 151–164 (1957).

[124] Ellis, R. *Bothalia* **12**(4), 641–671 (1979).

[125] Arogundade, O. O. and Adedeji, O. *Notulae Scientia Biologicae* **9**(1) (2017).

[126] Korn, R. W. *New Phytologist* **77**(1), 153–161 (1976).

[127] Stoddard, E. M. *Connecticut Agricultural Experimental Station Circular* **227** (1965).

[128] Vőfély, R. V., Pisano, G. D., Bartlett, M., and Braybrook, S. A. *under review* (2017).

[129] Sommer, C., Straehle, C., Koethe, U., and Hamprecht, F. A. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, 230–233. IEEE, (2011).

[130] Chase, M., Christenhusz, M., Fay, M., Byng, J., Judd, W., Soltis, D., Mabberley, D., Sennikov, A., Soltis, P., and Stevens, P. *Botanical Journal of the Linnean Society* **181**(1), 1–20 (2016).

[131] Zanne, A. E., Tank, D. C., Cornwell, W. K., Eastman, J. M., Smith, S. A., FitzJohn, R. G., McGlinn, D. J., OMeara, B. C., Moles, A. T., Reich, P. B., et al. *Nature* **506**, 89–92 (2014).

[132] Webb, C. O. and Donoghue, M. J. *Molecular Ecology Notes* **5**(1), 181–183 (2005).

[133] Keck, F., Rimet, F., Bouchez, A., and Franc, A. *Ecology and Evolution* **6**(9), 2774–80 (2016).

[134] Yu, G., Smith, D. K., Zhu, H., Guan, Y., and Lam, T. T.-Y. *Methods in Ecology and Evolution* **8**(1), 28–36 (2017).

[135] Revell, L. J., Harmon, L. J., and Collar, D. C. *Systematic Biology* **57**(4), 591–601 (2008).

[136] Blomberg, S. P. and Garland, T. *Journal of Evolutionary Biology* **15**(6), 899–910.

[137] Moran, P. A. P. *Journal of the Royal Statistical Society. Series B (Methodological)* **10**(2), 243–251 (1948).

[138] Moran, P. A. P. *Biometrika* **37**(1/2), 17–23 (1950).

[139] Gittleman, J. L. and Kot, M. *Systematic Zoology* **39**(3), 227–241 (1990).

[140] Anselin, L. *Geographical Analysis* **27**(2), 93–115.

[141] Camerarius, J. and Thal, J. *Hortus medicus et philosophicus: in quo plurimarum stirpium breves descriptiones, novae icones ... continentur ... Item, Sylva Hercynia, sive Catalogus plantarum sponte nascentium in montibus et locis plerisque Hercyniae Sylvae quae respicit Saxoniam, conscriptus.* Francofurti ad Moenum, [s.n.], (1588).

[142] Rédei, G. P. *Annual review of genetics* **9**(1), 111–127 (1975).

[143] Kaul, S., Koo, H. L., Jenkins, J., Rizzo, M., Rooney, T., Tallon, L. J., Feldblyum, T., Nierman, W., Benito, M. I., Lin, X., et al. *Nature* **408**(6814), 796–815 (2000).

[144] Kover, P. X., Valdar, W., Trakalo, J., Scarcelli, N., Ehrenreich, I. M., Purugganan, M. D., Durrant, C., and Mott, R. *PLOS Genetics* **5**(7) (2009).

[145] www.arabidopsis.org. *The Arabidopsis Information Resource (TAIR)* (2017-08-20).

[146] Bouton, S., Leboeuf, E., Mouille, G., Leydecker, M.-T., Talbotec, J., Granier, F., Lahaye, M., Höfte, H., and Truong, H.-N. *The Plant Cell* **14**(10), 2577–2590 (2002).

[147] Tsonis, A. A. and Tsonis, P. A. *Perspectives in Biology and Medicine* **30**(3), 355–361 (1987).

[148] Barbier de Reuille, P., Routier-Kierzkowska, A.-L., Kierzkowski, D., Bassel, G. W., Schpbach, T., Tauriello, G., Bajpai, N., Strauss, S., Weber, A., Kiss, A., Burian, A., Hofhuis, H., Sapala, A., Lipowczan, M., Heimlicher, M. B., Robinson, S., Bayer, E. M., Basler, K., Koumoutsakos, P., Roeder, A. H., Aegerter-Wilmsen, T., Nakayama, N., Tsiantis, M., Hay, A., Kwiatkowska, D., Xenarios, I., Kuhlemeier, C., and Smith, R. S. *eLife* **4**, e05864 (2015).

[149] Sanchez-Corrales, Y. E., Hartley, M., van Rooij, J., Maree, S. F. M., and Grieneisen, V. *bioRxiv* (2017).

[150] Blodgett, K. B. and Langmuir, I. *Phys. Rev.* **51**, 964–982 Jun (1937).

[151] Xue, N., Li, X., Bertulli, C., Li, Z., Patharagulpong, A., Sadok, A., and Huang, Y. Y. S. *PLOS ONE* **9**(4), e93590 (2014).

[152] Sapala, A., Runions, A., Routier-Kierzkowska, A.-L., Das Gupta, M., Hong, L., Hofhuis, H., Verger, S., Mosca, G., Li, C.-B., Hay, A., Hamant, O., Roeder, A. H., Tsiantis, M., Prusinkiewicz, P., and Smith, R. S. *eLife* **7**, e32794 (2018).

[153] Smith, L. G. *Nature Reviews Molecular Cell Biology* **2**(1), 33–39 (2001).

Appendix A

Elastic properties of composites

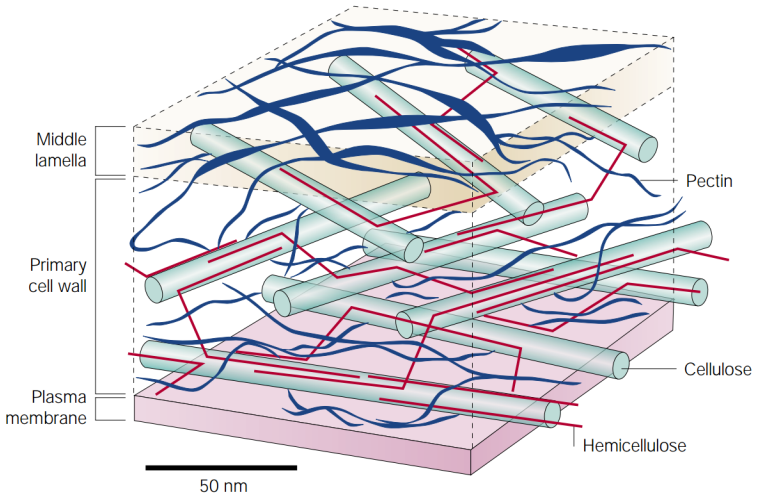


Figure A.1 – The structure of the cell wall
Image taken from [153].

The cell wall is a multi-component material: its main constituents are cellulose, pectin and hemicellulose — see Figure A.1. Cellulose forms bundles (microfibrils) that are embedded in the pectin matrix in layers parallel to the plasma membrane. The cellulose microfibrils are linked together by hemicellulose chains both within a given layer and across layers, thus creating a 3D network. This cellulosic network is embedded within a gel matrix

comprising pectins and structural proteins.

The general structure of the cell wall is very similar to that of fibre-reinforced composites. The simplest model is a two-component material where long, parallel fibres are embedded in a matrix. For simplicity we assume that both materials are in their elastic regime and described by the following equation: $\sigma = E\epsilon$, where σ denotes stress, ϵ denotes strain and Young's modulus, E, links the two quantity. To calculate the overall elastic modulus in this model, we need to consider the directionality.

1. When stretching the material with a force parallel to the fibres, we expect that the strain on both components is the same:

$$\epsilon_{\mathrm{f}} = \frac{\sigma_{\mathrm{f}}}{E_{\mathrm{f}}} = \frac{\sigma_{\mathrm{m}}}{E_{\mathrm{m}}} = \epsilon_{\mathrm{m}},$$

where subscripts f and m denote quantities relevant to the fibre and the matrix respectively.
Stress is, by definition, the force, F, per area, A: the total stress on the system is calculated as:

$$\sigma_{\mathrm{total}} = \frac{F_{\mathrm{total}}}{A_{\mathrm{total}}} = \frac{F_{\mathrm{f}} + F_{\mathrm{m}}}{A_{\mathrm{total}}} = \frac{\sigma_{\mathrm{f}} A_{\mathrm{f}} + \sigma_{\mathrm{m}} A_{\mathrm{m}}}{A_{\mathrm{total}}} = f\sigma_{\mathrm{f}} + (1-f)\sigma_{\mathrm{m}},$$

where f is the quantity known as the filling factor: this is the proportion of volume taken up by the fibres - which, in this arrangement is also the proportion of area at the surface perpendicular to the fibres.
Assuming that the strain on both components is the same, the resulting overall modulus is

$$E_{\mathrm{para}} = fE_{\mathrm{f}} + (1-f)E_{\mathrm{m}}.$$

2. Stretching this material perpendicular to the fibre orientation requires different constraints: in this configuration we assume that the stress is the same on both elements: $\sigma_{\mathrm{total}} = \sigma_{\mathrm{f}} = \sigma_{\mathrm{m}}$. The total strain is the sum of the strains of the two components: $\epsilon_{\mathrm{total}} = f\epsilon_{\mathrm{f}} + (1-f)\epsilon_{\mathrm{m}}$, again employing the filling factor, f. The total stress is then

$$\sigma_{\mathrm{total}} = E_{\mathrm{perp}}\epsilon_{\mathrm{total}} = E_{\mathrm{perp}}(f\epsilon_{\mathrm{f}} + (1-f)\epsilon_{\mathrm{m}}) = E_{\mathrm{perp}}(f\frac{\sigma_{\mathrm{f}}}{E_{\mathrm{f}}} + (1-f)\frac{\sigma_{\mathrm{m}}}{E_{\mathrm{m}}})$$

Using that the stress on the components are equal under perpendicular loading, we obtain the following relationship for the perpedicular modulus, E_{perp}:

$$\frac{1}{E_{\text{perp}}} = \frac{f}{E_{\text{f}}} + \frac{1-f}{E_{\text{m}}}$$

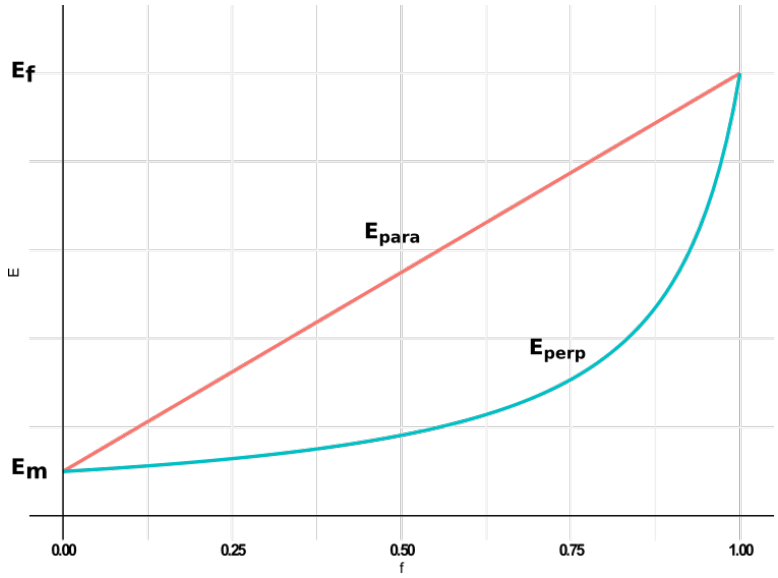


Figure A.2 – Elastic moduli parallel and perpendicular to the fibre orientation as a function of the filling factor. The parallel modulus is always greater than the perpendicular, therefore the preferred growth direction is perpendicular to the fibres.

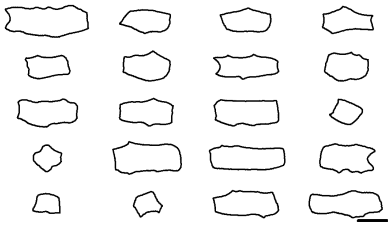
We see from Figure A.2 that the elastic modulus under parallel loading is always greater than under perpendicular loading. In practice this means that when cellulose fibres are aligned with each other, the preferential growth direction will be perpendicular to the fibres — it requires less energy to expand in this direction meaning that isotropic growth is no longer favoured. Because there is a globally defined preferential growth direction, the resulting growth is called anisotropic growth [1].

However, the cell wall can contain layers in which the cellulose fibres are oriented differently. Moreover, the orientation of cellulose fibres can differ between tissues at the same location. Cellulose orientation in a single cell therefore does not tell us which way an organ (or even a tissue) will expand.

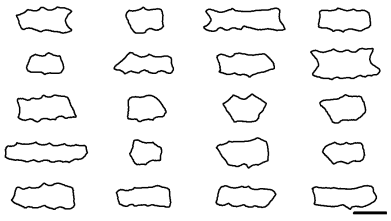
Appendix B

Cell shape data for the maize experiment

Here the segmented cell outlines are presented for every sample. The scale bar in these images is 100 µm.

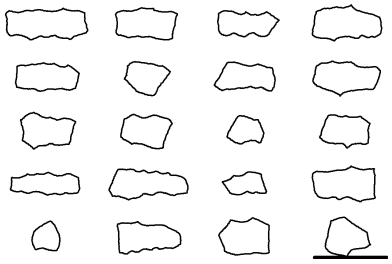


Week 1, leaf 1,0.5 cm

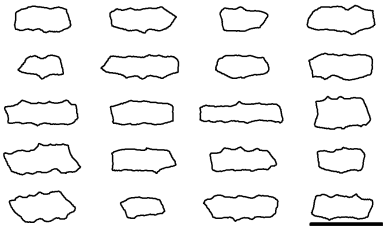


Week 1, leaf 1,1.5 cm

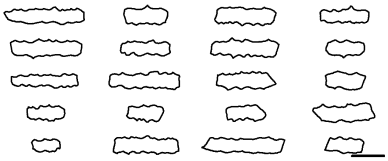
Week 1, leaf 1,4.5 cm



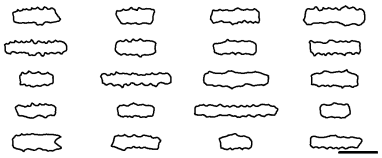
Week 1.5, leaf 1,0.5 cm



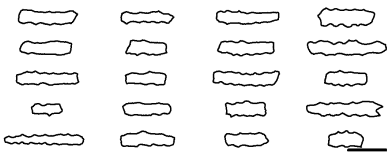
Week 1.5, leaf 1,2.5 cm



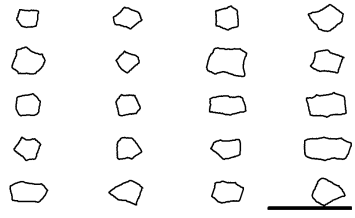
Week 1.5, leaf 1,3.5 cm



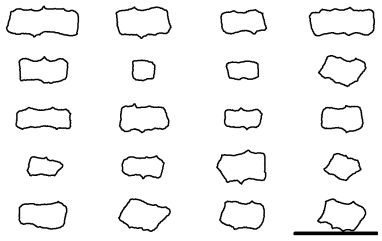
Week 1.5, leaf 1,4.5 cm



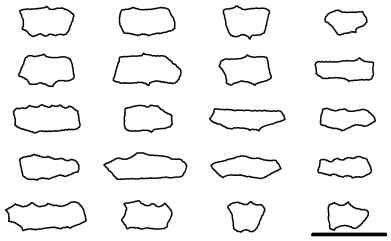
Week 1.5, leaf 2,0.5 cm



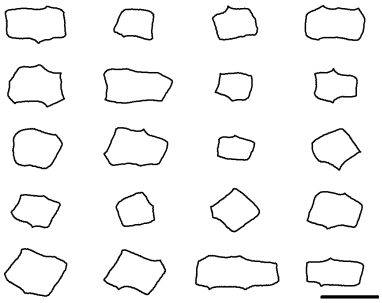
Week 1.5, leaf 2,2.5 cm



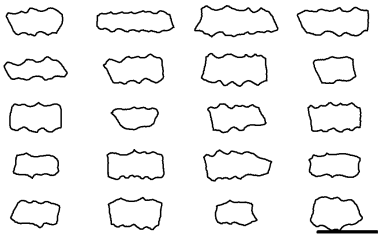
Week 1.5, leaf 2,4.5 cm



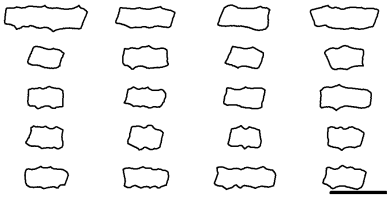
Week 1.5, leaf 2,11.5 cm



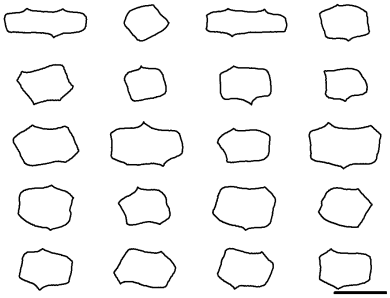
Week 2, leaf 1,0.5 cm



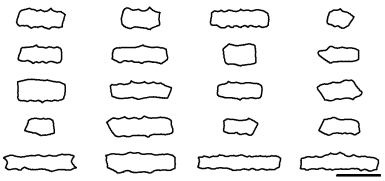
Week 2, leaf 1,2.5 cm



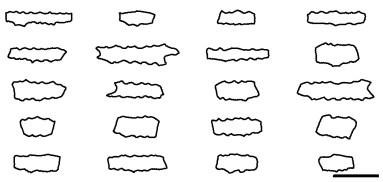
Week 2, leaf 1,4.5 cm



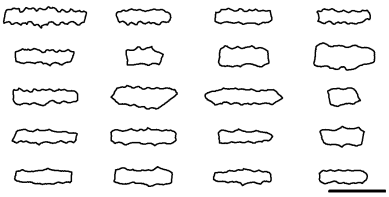
Week 2, leaf 2,0.5 cm



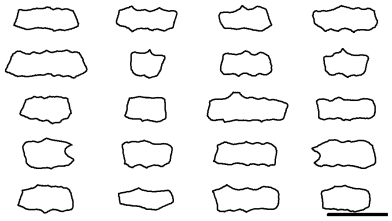
Week 2, leaf 2,5.5 cm



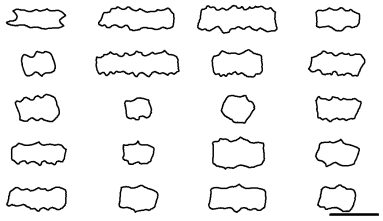
Week 2, leaf 2,10.5 cm



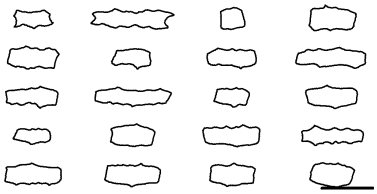
Week 2, leaf 2,14.5 cm



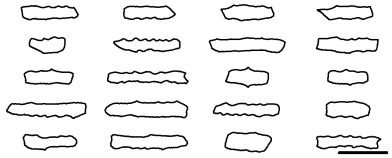
Week 2, leaf 3,0.5 cm



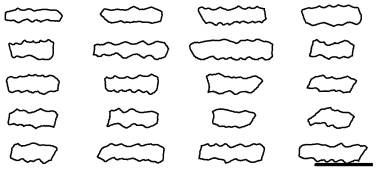
Week 2, leaf 3,12.5 cm



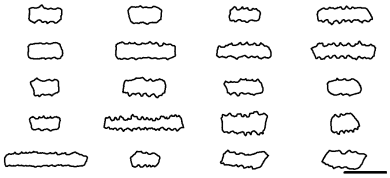
Week 2, leaf 3,18.5 cm



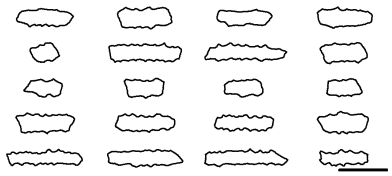
Week 2, leaf 3,23 cm



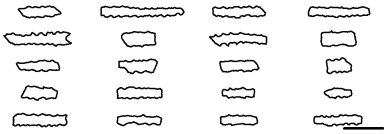
Week 2, leaf 3,27.5 cm



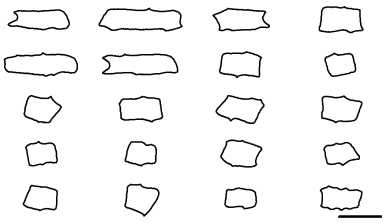
Week 2, leaf 4,0.5 cm



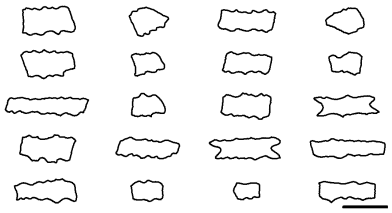
Week 2, leaf 4,9.5 cm



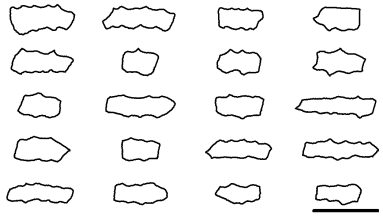
Week 2, leaf 4,18.5 cm



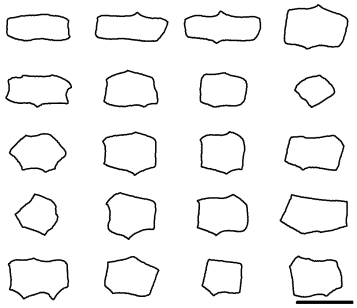
Week 2.5, leaf 1,0.5 cm



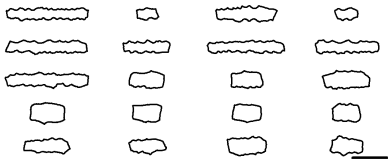
Week 2.5, leaf 1,2.5 cm



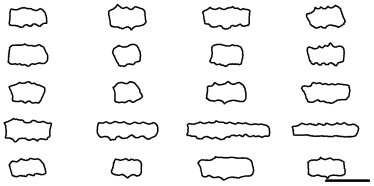
Week 2.5, leaf 1,4.5 cm



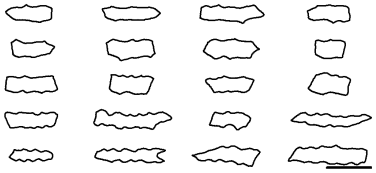
Week 2.5, leaf 2,0.5 cm



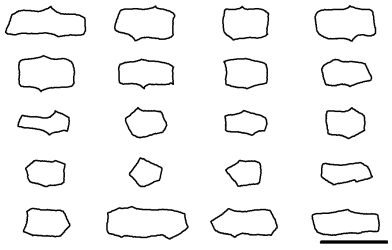
Week 2.5, leaf 2,4.5 cm



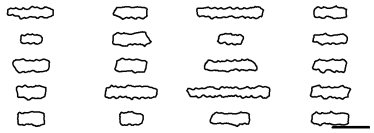
Week 2.5, leaf 2,8.5 cm



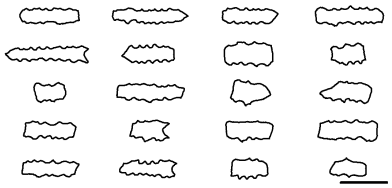
Week 2.5, leaf 2,12.5 cm



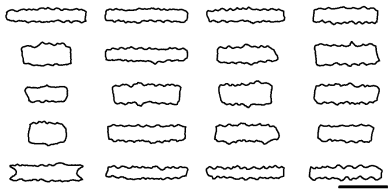
Week 2.5, leaf 3,0.5 cm



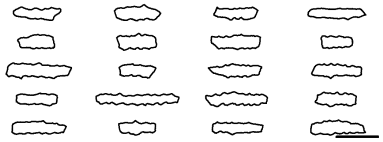
Week 2.5, leaf 3,5.5 cm



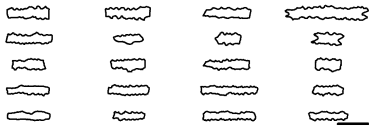
Week 2.5, leaf 3,17.5 cm



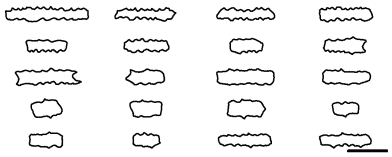
Week 2.5, leaf 3,25.5 cm



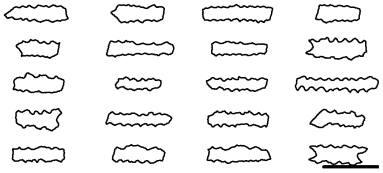
Week 2.5, leaf 4,0.5 cm



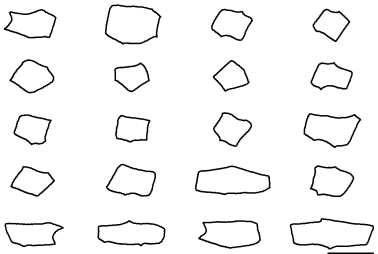
Week 2.5, leaf 4,7.5 cm



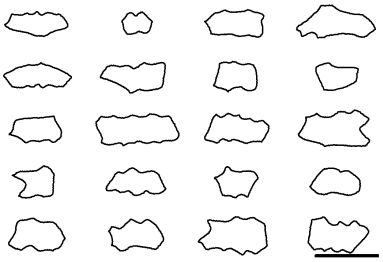
Week 2.5, leaf 4,14.5 cm



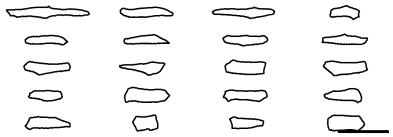
Week 2.5, leaf 4,21.5 cm



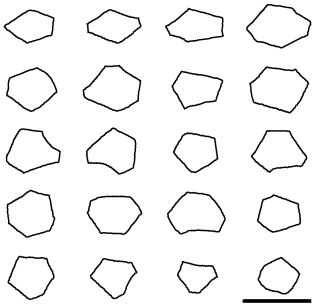
Week 3, leaf 1,0.5 cm



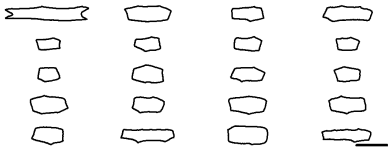
Week 3, leaf 1,2 cm



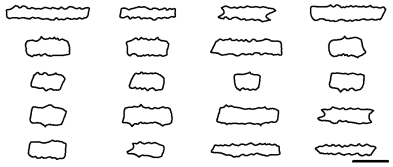
Week 3, leaf 1,3.5 cm



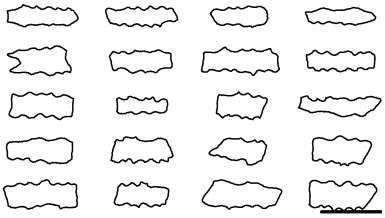
Week 3, leaf 2,0.5 cm



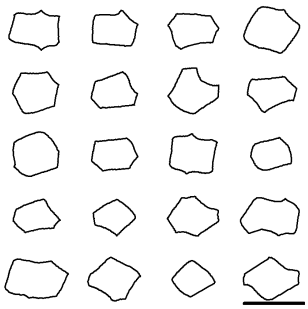
Week 3, leaf 2,4.5 cm



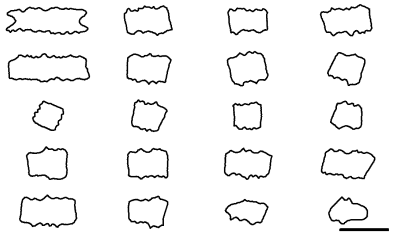
Week 3, leaf 2,8.5 cm



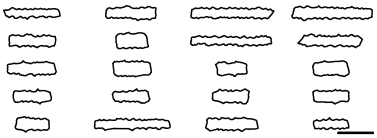
Week 3, leaf 2,12.5 cm



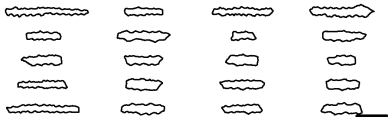
Week 3, leaf 3,0.5 cm



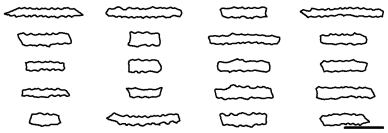
Week 3, leaf 3,6.5 cm



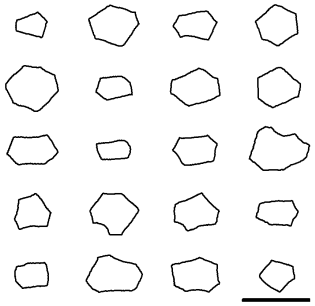
Week 3, leaf 3,12.5 cm



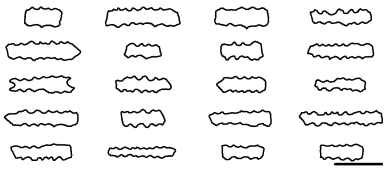
Week 3, leaf 3,18.5 cm



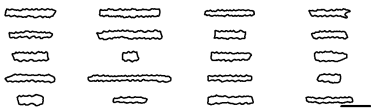
Week 3, leaf 3,24.5 cm



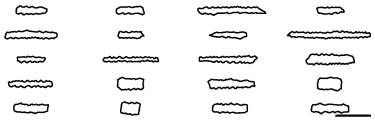
Week 3, leaf 4,0.5 cm



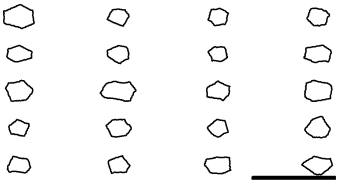
Week 3, leaf 4,11.5 cm



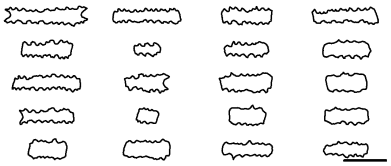
Week 3, leaf 4,22.5 cm



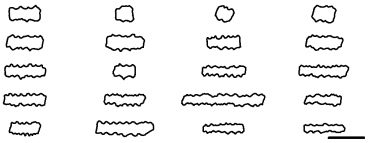
Week 3, leaf 4,35.5 cm



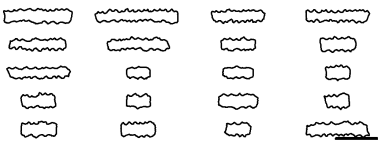
Week 3, leaf 5,0.5 cm



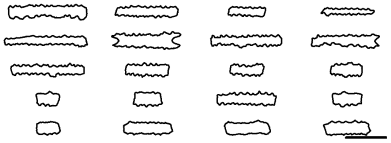
Week 3, leaf 5,11.5 cm



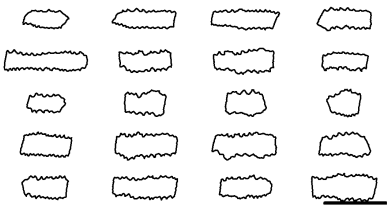
Week 3, leaf 5,22.5 cm



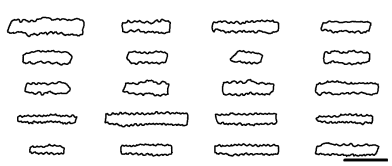
Week 3, leaf 5,33.5 cm



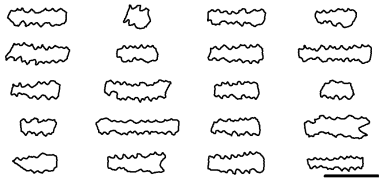
Week 3, leaf 5,44.5 cm



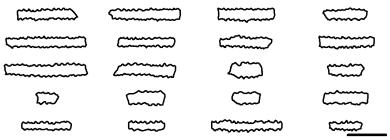
Week 3, leaf 6,0.5 cm



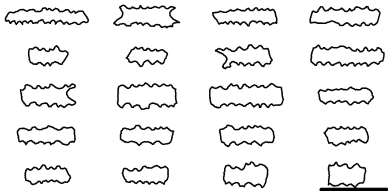
Week 3, leaf 6,11.5 cm



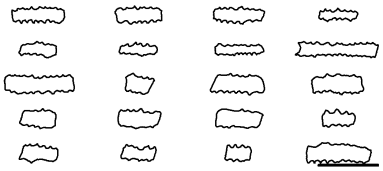
Week 3, leaf 6,22.5 cm



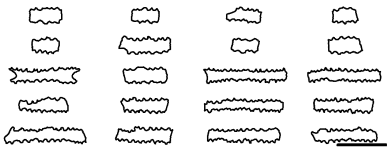
Week 3, leaf 6,33.5 cm



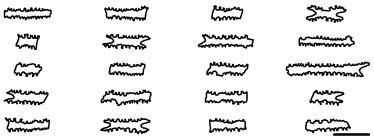
Week 3, leaf 6,44.5 cm



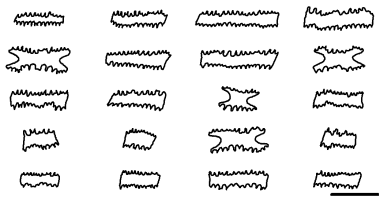
Week 3, leaf 7,0.5 cm



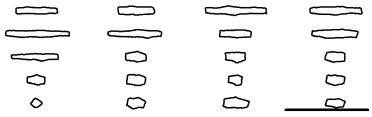
Week 3, leaf 7,11.5 cm



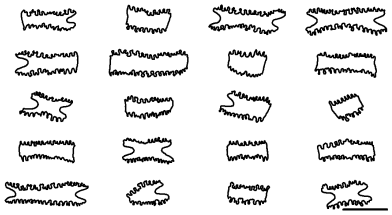
Week 3, leaf 7,22.5 cm



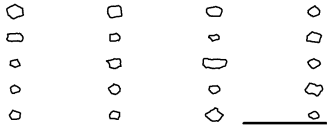
Week 3, leaf 7,33.5 cm



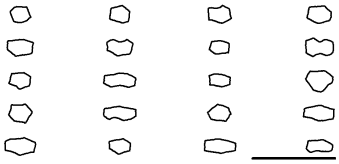
Week 4, leaf 7,0.5 cm



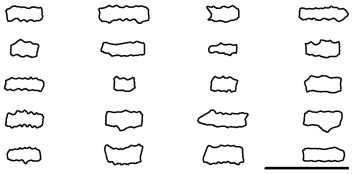
Week 4, leaf 7,8.5 cm



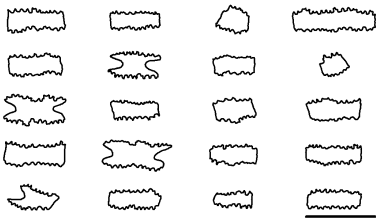
Week 4, leaf 8,0.5 cm



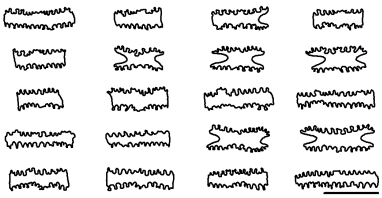
Week 4, leaf 8,1.5 cm



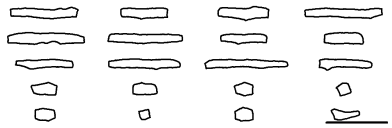
Week 4, leaf 8,2.5 cm



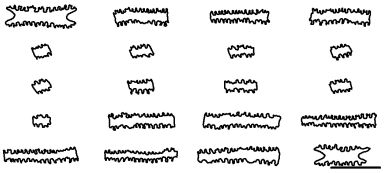
Week 4, leaf 8,4.5 cm



Week 4, leaf 8,7.5 cm

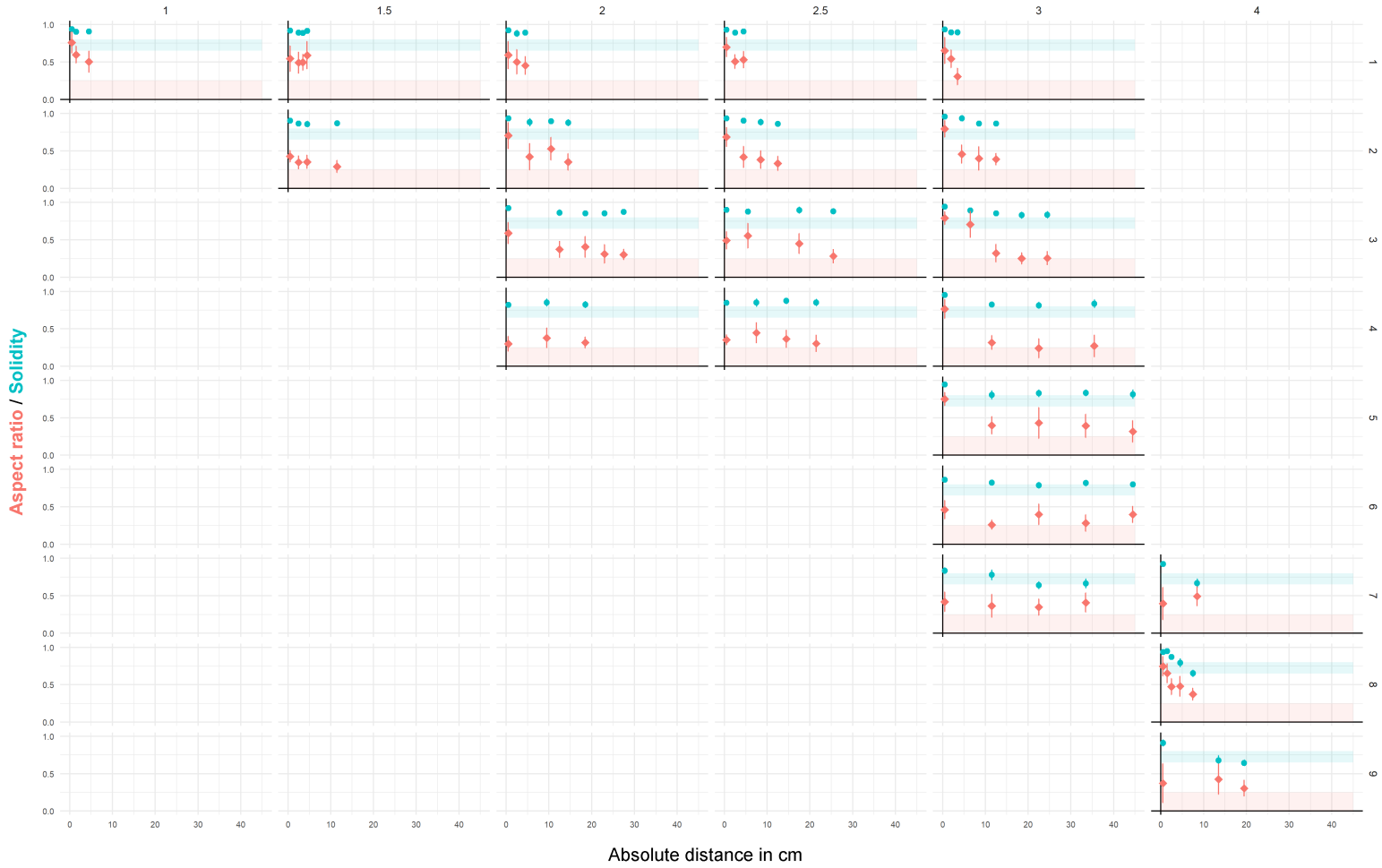


Week 4, leaf 9,0.5 cm

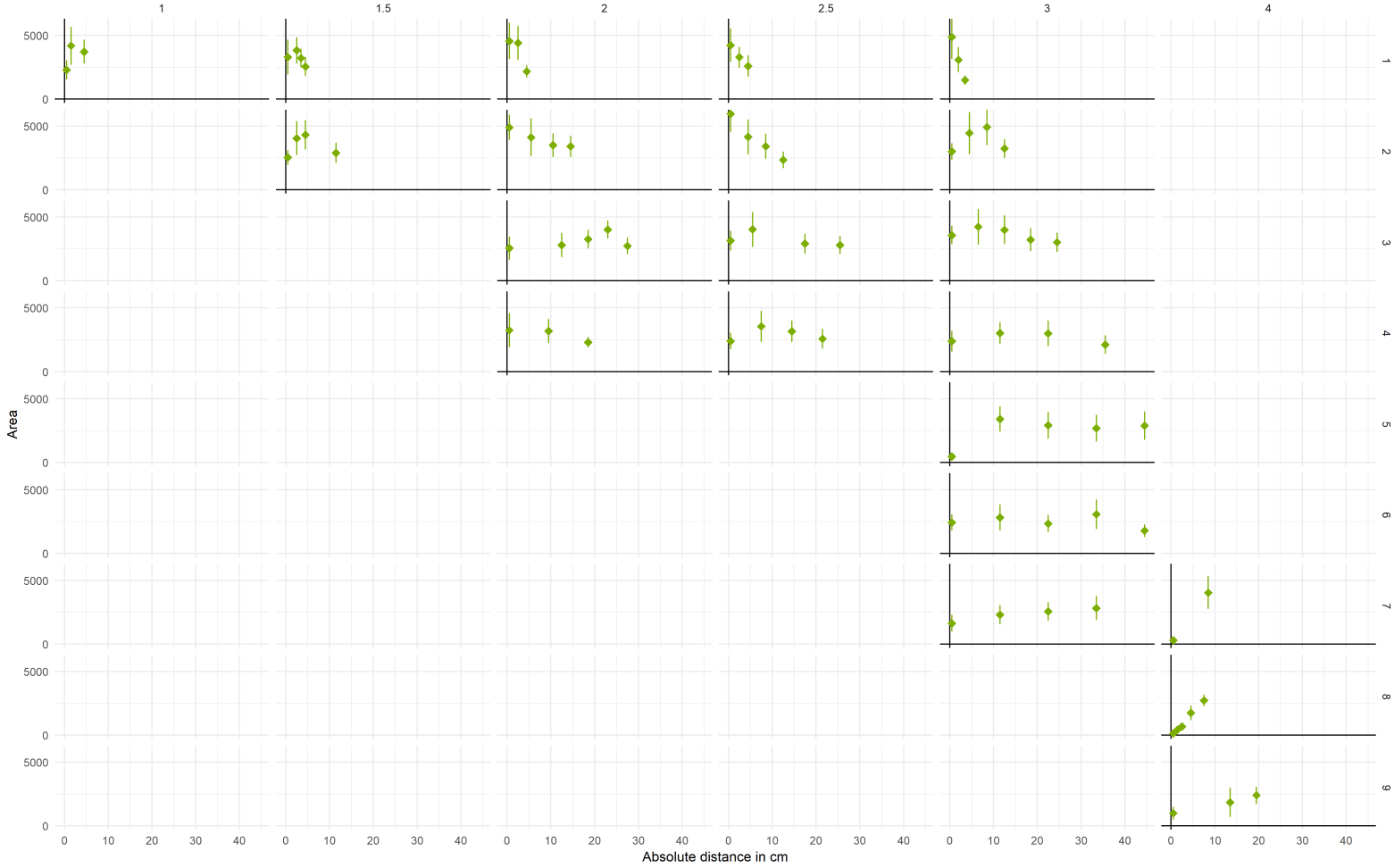


Week 4, leaf 9,13.5 cm

Week 4, leaf 9,19.5 cm



[
Variation of the cell shape along the maize leaf]**Variation of the cell shape along the maize leaf**
Average elongation is expressed with the aspect ratio (red diamonds), average undulation depth is represented by solidity (blue circles). Absolute distance is measured in cm from the ligule. Columns correspond to the age of the seedling in weeks (1, 1.5, 2, 2.5, 3, 4), rows correspond to leaf number (outermost = 1). Circles mark data from individual cells, diamond marks the sample average.
Coloured regions are added to guide the eye: the blue region indicates the difference between minimum average solidity values of juvenile and adult leaves, the red region covers aspect ratio values smaller than the minimal sample average observed.



[

Variation of the cell area along the maize leaf]**Variation of the cell area along the maize leaf**
Average area is measured in µm^2, distance is measured in cm from the ligule. Columns correspond to the age of the seedling in weeks (1, 1.5, 2, 2.5, 3, 4), rows correspond to leaf number (outermost = 1).

Appendix C

Cell shape data for vascular plants

The following tables contain information about the species used in this analysis. The first one provides taxonomical information: name, family and order; as well as the location it was collected from and which of the five groups it was placed for analysis. The second one presents the resulting cell shape and leaf shape quantifiers for these species, separately for the abaxial and adaxial sides.

species	group	order	family	origin
Osmunda banksiifolia	Ferns	Osmundales	Osmundaceae	Amherst
Stenochlaena palustris	Ferns	Polypodiales	Blechnaceae	Amherst
Nephrolepis sp.	Ferns	Polypodiales	Davalliaceae	Amherst
Phlebodium aureum	Ferns	Polypodiales	Polypodiaceae	Amherst
Cycas revoluta	Gymnosperms	Cycadales	Cycadaceae	Amherst
Bowenia serrulata	Gymnosperms	Cycadales	Zamiaceae	Amherst
Encephalartos transvenosus	Gymnosperms	Cycadales	Zamiaceae	Amherst
Microcycas calocoma	Gymnosperms	Cycadales	Zamiaceae	Amherst
Zamia furfuraceae	Gymnosperms	Cycadales	Zamiaceae	Amherst
Zamia pumila	Gymnosperms	Cycadales	Zamiaceae	Amherst
Zamia skinneri	Gymnosperms	Cycadales	Zamiaceae	Amherst
Ginkgo biloba	Gymnosperms	Ginkgoales	Ginkgoaceae	Amherst
Gnetum montanum	Gymnosperms	Gnetales	Gnetaceae	Amherst
Picea abies	Gymnosperms	Pinales	Pinaceae	Amherst
Picea glauca	Gymnosperms	Pinales	Pinaceae	Amherst
Tsuga canadensis	Gymnosperms	Pinales	Pinaceae	Amherst
Araucaria sp.	Gymnosperms	Araucariales	Araucariaceae	Amherst
Araucaria bidwillii	Gymnosperms	Araucariales	Araucariaceae	Amherst
Podocarpus macrophyllus	Gymnosperms	Araucariales	Podocarpaceae	Amherst
Taxodium distichum	Gymnosperms	Cupressales	Cupressaceae	Amherst
Widdringtonia nodiflora	Gymnosperms	Cupressales	Cupressaceae	Amherst
Chamaecyparis thyoides	Gymnosperms	Cupressales	Cupressaceae	Amherst
Taxus canadensis	Gymnosperms	Cupressales	Taxaceae	Amherst
Illicium anisatum	Early diverging angiosperms	Austrobaileyales	Schisandraceae	Amherst
Annona montana	Early diverging angiosperms	Magnoliales	Annonaceae	Amherst
Laurus nobilis	Early diverging angiosperms	Laurales	Lauraceae	Amherst
Asarum europaeum	Early diverging angiosperms	Piperales	Aristolochiaceae	Cambridge
Peperomia sp.	Early diverging angiosperms	Piperales	Piperaceae	Amherst
Drimys winteri	Early diverging angiosperms	Canellales	Winteraceae	Amherst

species	group	order	family	origin
Chloranthus sp.	Early diverging angiosperms	Chloranthales	Chloranthaceae	Amherst
Sabal minor	Monocots	Arecales	Arecaceae	Amherst
Carex morrowii	Monocots	Poales	Cyperaceae	Amherst
Avena strigosa	Monocots	Poales	Poaceae	Amherst
Danthonia californica	Monocots	Poales	Poaceae	Amherst
Eriachne sp.	Monocots	Poales	Poaceae	Amherst
Hordeum vulgare	Monocots	Poales	Poaceae	Amherst
Oryza sativa	Monocots	Poales	Poaceae	Amherst
Phalaris arundinacea	Monocots	Poales	Poaceae	Cambridge
Setaria italica	Monocots	Poales	Poaceae	Amherst
Sorghum bicolor	Monocots	Poales	Poaceae	Amherst
Triraphis sp.	Monocots	Poales	Poaceae	Amherst
Triticum aestivum	Monocots	Poales	Poaceae	Cambridge
Palisota mannii	Monocots	Commelinales	Commelinaceae	Cambridge
Pollia japonica	Monocots	Commelinales	Commelinaceae	Cambridge
Tradescantia spathacea sw. 'Vittata'	Monocots	Commelinales	Commelinaceae	Cambridge
Tradescantia virginiana	Monocots	Commelinales	Commelinaceae	Cambridge
Eichhornia crassipes	Monocots	Commelinales	Pontederiaceae	Amherst
Costus igneus	Monocots	Zingiberales	Costaceae	Cambridge
Monocostus uniflorus	Monocots	Zingiberales	Costaceae	Cambridge
Alpinia zerumbet	Monocots	Zingiberales	Zingiberaceae	Amherst
Curcuma longa	Monocots	Zingiberales	Zingiberaceae	Cambridge
Zingiber officinale Roscoe	Monocots	Zingiberales	Zingiberaceae	Cambridge
Cirrhopetalum gracillimum	Monocots	Asparagales	Orchidaceae	Cambridge
Cymbidium sp.	Monocots	Asparagales	Orchidaceae	Amherst
Oncidium sphacelatum	Monocots	Asparagales	Orchidaceae	Cambridge
Vanilla planifolia	Monocots	Asparagales	Orchidaceae	Cambridge
Crocus sativus	Monocots	Asparagales	Iridaceae	Cambridge
Iris germanica 'Kharput'	Monocots	Asparagales	Iridaceae	Cambridge

species	group	order	family	origin
Iris caucasica 'Rainbow Grand'	Monocots	Asparagales	Iridaceae	Cambridge
Aloe vera	Monocots	Asparagales	Asphodelaceae	Amherst
Haworthia retusa	Monocots	Asparagales	Asphodelaceae	Amherst
Hemerocallis fulva	Monocots	Asparagales	Asphodelaceae	Cambridge
Kniphofia caulescens	Monocots	Asparagales	Asphodelaceae	Cambridge
Agapanthus preacox	Monocots	Asparagales	Amaryllidaceae	Amherst
Amaryllis belladonna	Monocots	Asparagales	Amaryllidaceae	Cambridge
Clivia miniata	Monocots	Asparagales	Amaryllidaceae	Amherst
Leucojum aestivum	Monocots	Asparagales	Amaryllidaceae	Cambridge
Sternbergia lutea	Monocots	Asparagales	Amaryllidaceae	Cambridge
Agave sisalana	Monocots	Asparagales	Asparagaceae	Amherst
Calibanus hookeri	Monocots	Asparagales	Asparagaceae	Amherst
Danae racemosa	Monocots	Asparagales	Asparagaceae	Cambridge
Ornithogalum caudatum	Monocots	Asparagales	Asparagaceae	Amherst
Rohdea japonica	Monocots	Asparagales	Asparagaceae	Cambridge
Ruscus aculeatus	Monocots	Asparagales	Asparagaceae	Cambridge
Ruscus hypoglossum	Monocots	Asparagales	Asparagaceae	Amherst
Yucca gloriosa	Monocots	Asparagales	Asparagaceae	Cambridge
Alstroemeria aurea	Monocots	Liliales	Alstroemeriaceae	Cambridge
Luzuriaga latifolia	Monocots	Liliales	Alstroemeriaceae	Amherst
Smilax bona-nox 'Cantab'	Monocots	Liliales	Smilacaceae	Cambridge
Tricyrtis hirta	Monocots	Liliales	Liliaceae	Cambridge
Dioscorea bulbifera	Monocots	Dioscoreales	Dioscoreaceae	Amherst
Carludovica palmata	Monocots	Pandanales	Cyclanthaceae	Amherst
Freycinetia sp.	Monocots	Pandanales	Pandanaceae	Amherst
Pandanus veitchei	Monocots	Pandanales	Pandanaceae	Amherst
Stichoneuron caudatum	Monocots	Pandanales	Stemonaceae	Amherst
Aglaonema crispum	Monocots	Alismatales	Araceae	Cambridge
Amorphaphallus titanum	Monocots	Alismatales	Araceae	Amherst

species	group	order	family	origin
Arum maculatum	Monocots	Alismatales	Araceae	Cambridge
Dracunculus canariensis	Monocots	Alismatales	Araceae	Cambridge
Rhektophyllum mirabile	Monocots	Alismatales	Araceae	Cambridge
Spathiphyllum wallisii	Monocots	Alismatales	Araceae	Cambridge
Zantedeschia aethiopica	Monocots	Alismatales	Araceae	Cambridge
Hydrocleys nymphoides	Monocots	Alismatales	Alismataceae	Amherst
Acorus gramineus	Monocots	Acorales	Acoraceae	Amherst
Aconitum carmichaelii var. arendsi	Eudicots	Ranunculales	Ranunculaceae	Cambridge
Anemone canadensis	Eudicots	Ranunculales	Ranunculaceae	Cambridge
Helleborus orientalis ssp orientalis	Eudicots	Ranunculales	Ranunculaceae	Cambridge
Astragalus falcatus	Eudicots	Fabales	Fabaceae	Cambridge
Dorycinum rectum	Eudicots	Fabales	Fabaceae	Cambridge
Medicago sativa	Eudicots	Fabales	Fabaceae	Cambridge
Mimosa pudica	Eudicots	Fabales	Fabaceae	Amherst
Senna didymobotrya	Eudicots	Fabales	Fabaceae	Amherst
Trifolium pannonicum	Eudicots	Fabales	Fabaceae	Cambridge
Ulex europaeus	Eudicots	Fabales	Fabaceae	Cambridge
Geum triflorum 'Purple Avens'	Eudicots	Rosales	Rosaceae	Cambridge
Rosa damascena 'versicolor'	Eudicots	Rosales	Rosaceae	Cambridge
Lithocarpus henryi	Eudicots	Fagales	Fagaceae	Amherst
Oxalis regnellii	Eudicots	Oxalidales	Oxalidaceae	Amherst
Oxalis tetraphylla	Eudicots	Oxalidales	Oxalidaceae	Cambridge
Oxalis valdiviensis	Eudicots	Oxalidales	Oxalidaceae	Cambridge
Hiptage benghalensis	Eudicots	Malpighiales	Malpighiaceae	Amherst
Passiflora edulis	Eudicots	Malpighiales	Passifloraceae	Amherst
Passiflora sp.	Eudicots	Malpighiales	Passifloraceae	Amherst
Poliothyrsis sinensis	Eudicots	Malpighiales	Salicaceae	Cambridge
Euphorbia flanaganii	Eudicots	Malpighiales	Euphorbiaceae	Amherst
Euphorbia mellifera	Eudicots	Malpighiales	Euphorbiaceae	Cambridge

species	group	order	family	origin
Euphorbia pulcherrima	Eudicots	Malpighiales	Euphorbiaceae	Amherst
Catha edulis	Eudicots	Celastrales	Celastraceae	Amherst
Guaiacum officinale	Eudicots	Zygophyllales	Zygophyllaceae	Amherst
Pelargonium carnosum	Eudicots	Geraniales	Geraniaceae	Amherst
Cuphea ignea	Eudicots	Myrtales	Lythraceae	Cambridge
Heimia myrthifolia	Eudicots	Myrtales	Lythraceae	Cambridge
Fuchsia magellanica var. molinae	Eudicots	Myrtales	Onagraceae	Cambridge
Fuchsia 'Mrs Popple'	Eudicots	Myrtales	Onagraceae	Cambridge
Oenothera stricta	Eudicots	Myrtales	Onagraceae	Cambridge
Ceiba pentandra	Eudicots	Malvales	Malvaceae	Amherst
Malva sylvestris	Eudicots	Malvales	Malvaceae	Cambridge
Carica papaya	Eudicots	Brassicales	Caricaceae	Amherst
Erysimum scoparium	Eudicots	Brassicales	Brassicaceae	Cambridge
Sisymbrium austriaticum	Eudicots	Brassicales	Brassicaceae	Cambridge
Bursera schlechtendalii	Eudicots	Sapindales	Burseraceae	Amherst
Rhus potaninii	Eudicots	Sapindales	Anacardiaceae	Cambridge
Dodonaea viscosa	Eudicots	Sapindales	Sapindaceae	Amherst
Citrus x limon	Eudicots	Sapindales	Rutaceae	Cambridge
Murraya koenigii	Eudicots	Sapindales	Rutaceae	Amherst
Cissus quadrangularis	Eudicots	Vitales	Vitaceae	Amherst
Cissus tuberosa	Eudicots	Vitales	Vitaceae	Amherst
Paeonia tenuifolia	Eudicots	Saxifragales	Paeoniaceae	Cambridge
Bergenia purpurascens	Eudicots	Saxifragales	Saxifragaceae	Cambridge
Saxifraga canaliculata	Eudicots	Saxifragales	Saxifragaceae	Cambridge
Saxifraga hostii ssp hostii	Eudicots	Saxifragales	Saxifragaceae	Cambridge
Adromischus sp.	Eudicots	Saxifragales	Crassulaceae	Amherst
Crassula fasciatudform	Eudicots	Saxifragales	Crassulaceae	Amherst
Haloragis erecta	Eudicots	Saxifragales	Haloragaceae	Cambridge
Ceratostigma plumbaginoides	Eudicots	Caryophyllales	Plumbaginaceae	Cambridge

species	group	order	family	origin
Ceratostigma wilmottianum	Eudicots	Caryophyllales	Plumbaginaceae	Cambridge
Homalocladium platycladum	Eudicots	Caryophyllales	Polygonaceae	Amherst
Oxyria digyna	Eudicots	Caryophyllales	Polygonaceae	Cambridge
Persicaria affinis	Eudicots	Caryophyllales	Polygonaceae	Cambridge
Persicaria polystachya	Eudicots	Caryophyllales	Polygonaceae	Cambridge
Persicaria weyrichii	Eudicots	Caryophyllales	Polygonaceae	Cambridge
Rumex acetosella	Eudicots	Caryophyllales	Polygonaceae	Cambridge
Rumex scutatus	Eudicots	Caryophyllales	Polygonaceae	Cambridge
Amaranthus caudatus	Eudicots	Caryophyllales	Amaranthaceae	Cambridge
Amaranthus hybridus	Eudicots	Caryophyllales	Amaranthaceae	Cambridge
Beta trigyna	Eudicots	Caryophyllales	Amaranthaceae	Cambridge
Beta vulgaris ssp. vulgaris	Eudicots	Caryophyllales	Amaranthaceae	Cambridge
Chenopodium bonus-henricus	Eudicots	Caryophyllales	Amaranthaceae	Cambridge
Ercilla volubilis	Eudicots	Caryophyllales	Phytolaccaceae	Cambridge
Epiphyllum sp.	Eudicots	Caryophyllales	Cactaceae	Amherst
Impatiens repens	Eudicots	Ericales	Balsaminaceae	Amherst
Ardisia crispa	Eudicots	Ericales	Primulaceae	Amherst
Macleania insignis	Eudicots	Ericales	Ericaceae	Amherst
Ilex paraguariensis	Eudicots	Aquifoliales	Aquifoliaceae	Amherst
Campanula fenestrellata	Eudicots	Asterales	Campanulaceae	Cambridge
Campanula poscharskyana	Eudicots	Asterales	Campanulaceae	Cambridge
Jasione heldreichii	Eudicots	Asterales	Campanulaceae	Cambridge
Berkheya purpurea	Eudicots	Asterales	Asteraceae	Cambridge
Berkheya radula	Eudicots	Asterales	Asteraceae	Cambridge
Centaurea dealbata	Eudicots	Asterales	Asteraceae	Cambridge
Centaurea simplicicaulis	Eudicots	Asterales	Asteraceae	Cambridge
Othonna crassifolia	Eudicots	Asterales	Asteraceae	Amherst
Picris echioides	Eudicots	Asterales	Asteraceae	Cambridge
Scorzonera hispanica	Eudicots	Asterales	Asteraceae	Cambridge

species	group	order	family	origin
Hedera nepalensis	Eudicots	Apiales	Araliaceae	Amherst
Bupleurum fruticosum	Eudicots	Apiales	Apiaceae	Cambridge
Eryngium agavifolium	Eudicots	Apiales	Apiaceae	Cambridge
Eryngium bourgatii	Eudicots	Apiales	Apiaceae	Cambridge
Myrrhis odorata	Eudicots	Apiales	Apiaceae	Cambridge
Pastinaca sativa 'Tender and true'	Eudicots	Apiales	Apiaceae	Cambridge
Centrathus ruber	Eudicots	Dipsacales	Caprifoliaceae	Cambridge
Cephalaria flava	Eudicots	Dipsacales	Caprifoliaceae	Cambridge
Lonicera quinquelocularis	Eudicots	Dipsacales	Caprifoliaceae	Cambridge
Scabiosa olgae	Eudicots	Dipsacales	Caprifoliaceae	Cambridge
Succisella inflexa	Eudicots	Dipsacales	Caprifoliaceae	Cambridge
Valeriana phu	Eudicots	Dipsacales	Caprifoliaceae	Cambridge
Brugmansia suaveolens	Eudicots	Solanales	Solanaceae	Amherst
Capsicum annuum 'Etna'	Eudicots	Solanales	Solanaceae	Cambridge
Capsicum annuum 'Mohawk'	Eudicots	Solanales	Solanaceae	Cambridge
Salpichroa origanifolia	Eudicots	Solanales	Solanaceae	Cambridge
Montinia caryophyllacea	Eudicots	Solanales	Montiniaceae	Amherst
Jasminum fruticans	Eudicots	Lamiales	Oleaceae	Cambridge
Jasminum humile var. revolutum	Eudicots	Lamiales	Oleaceae	Cambridge
Syringa 'Charles X'	Eudicots	Lamiales	Oleaceae	Cambridge
Syringa vulgaris	Eudicots	Lamiales	Oleaceae	Cambridge
Globularia punctata	Eudicots	Lamiales	Plantaginaceae	Cambridge
Globularia trichosantha	Eudicots	Lamiales	Plantaginaceae	Cambridge
Hebe rakaiensis	Eudicots	Lamiales	Plantaginaceae	Cambridge
Penstemon 'Raven'	Eudicots	Lamiales	Plantaginaceae	Cambridge
Plantago cynops	Eudicots	Lamiales	Plantaginaceae	Cambridge
Synthyris missurica var stellata	Eudicots	Lamiales	Plantaginaceae	Cambridge
Veronica gentianoides	Eudicots	Lamiales	Plantaginaceae	Cambridge
Veronica petraea	Eudicots	Lamiales	Plantaginaceae	Cambridge

species	group	order	family	origin
Scrophularia canina ssp. Bicolor	Eudicots	Lamiales	Scrophulariaceae	Cambridge
Scrophularia heterophylla	Eudicots	Lamiales	Scrophulariaceae	Cambridge
Acanthus hungaricus	Eudicots	Lamiales	Acanthaceae	Cambridge
Acanthus spinosus	Eudicots	Lamiales	Acanthaceae	Cambridge
Justicia guttata	Eudicots	Lamiales	Acanthaceae	Amherst
Verbena bonariensis	Eudicots	Lamiales	Verbenaceae	Cambridge
Callicarpa japonica	Eudicots	Lamiales	Lamiaceae	Cambridge
Clerodendrum thomsoniae	Eudicots	Lamiales	Lamiaceae	Amherst
Lamium orvala	Eudicots	Lamiales	Lamiaceae	Cambridge
Mentha piperita	Eudicots	Lamiales	Lamiaceae	Cambridge
Salvia officinalis 'Purpurascens'	Eudicots	Lamiales	Lamiaceae	Cambridge
Stachys macrantha	Eudicots	Lamiales	Lamiaceae	Cambridge
Galium odoratum	Eudicots	Gentianales	Rubiaceae	Cambridge
Galium rubioides	Eudicots	Gentianales	Rubiaceae	Cambridge
Rubia tinctorum	Eudicots	Gentianales	Rubiaceae	Cambridge
Fagraea berteriana	Eudicots	Gentianales	Gentianaceae	Amherst
Catharanthus roseus	Eudicots	Gentianales	Apocynaceae	Amherst
Ceropegia sandersonii	Eudicots	Gentianales	Apocynaceae	Amherst
Rauvolfia verticillata	Eudicots	Gentianales	Apocynaceae	Amherst
Tabernaemontana divaricata	Eudicots	Gentianales	Apocynaceae	Amherst
Anchusa officinalis	Eudicots	Boraginales	Boraginaceae	Cambridge
Borago officinalis	Eudicots	Boraginales	Boraginaceae	Cambridge
Cerinthe minor	Eudicots	Boraginales	Boraginaceae	Cambridge
Cordia nitida	Eudicots	Boraginales	Boraginaceae	Amherst
Nemophila menziesii	Eudicots	Boraginales	Boraginaceae	Cambridge
Symphytum caucasium x S. orientale	Eudicots	Boraginales	Boraginaceae	Cambridge
Symphytum grandiflorum	Eudicots	Boraginales	Boraginaceae	Cambridge

species	side	cell area (µm^2)	aspect ratio	circularity	solidity	leaf area (mm^2)	aspect ratio	circularity	solidity
Osmunda banksiifolia	ad	5,461	0.666	7.73	0.551	45.8	0.203	2.99	0.889
Stenochlaena palustris	ad	9,525	0.363	5.65	0.623	29.4	0.378	1.96	0.961
Nephrolepis sp.	ad	5,586	0.552	2.76	0.731	4.4	0.406	1.56	0.990
Phlebodium aureum	ad	2,232	0.714	3.44	0.685	10.7	0.065	8.77	0.814
Cycas revoluta	ad	1,616	0.490	2.14	0.855	8.6	0.206	4.92	0.892
Bowenia serrulata	ad	1,216	0.191	4.55	0.776	74.6	0.425	1.57	0.988
Encephalartos transvenosus	ad	812	0.316	3.19	0.775	1.6	0.135	4.10	0.960
Microcycas calocoma	ad	1,353	0.303	3.12	0.789	17.5	0.217	4.36	0.841
Zamia furfuraceae	ad	1,656	0.168	4.85	0.776	12.5	0.214	2.79	0.944
Zamia pumila	ad	1,771	0.256	3.84	0.756	46.6	0.518	1.48	0.985
Zamia skinneri	ad	1,419	0.388	2.89	0.786	46.6	0.347	2.03	0.965
Ginkgo biloba	ad	4,130	0.563	2.29	0.815				
Gnetum montanum	ad	1,032	0.663	1.87	0.856	76.7	0.255	2.67	0.942
Picea abies	ad	2,975	0.289	3.07	0.797				
Picea glauca	ad	2,090	0.536	1.98	0.865				
Tsuga canadensis	ad	3,251	0.116	5.44	0.828				
Araucaria sp.	ad	2,054	0.139	4.70	0.837	21.7	0.123	10.38	0.506
Araucaria bidwillii	ad	1,143	0.251	2.93	0.864	0.9	0.097	4.51	0.953
Podocarpus macrophyllus	ad	296	0.294	3.11	0.792	23.7	0.613	2.11	0.946
Taxodium distichum	ad	3,052	0.422	2.38	0.823	0.2	0.133	4.90	0.745
Widdringtonia nodiflora	ad	4,984	0.137	4.34	0.872	9.6	0.044	11.51	0.895
Chamaecyparis thyoides	ad	449	0.358	2.41	0.828				
Taxus canadensis	ad	2,568	0.600	1.55	0.934	11.1	0.574	1.32	0.995
Illicium anisatum	ad	645	0.645	2.10	0.812	22.8	0.354	1.80	0.992
Annona montana	ad	2,593	0.727	1.39	0.934				
Laurus nobilis	ad	1,037	0.746	4.13	0.609	6.1	0.508	1.81	0.904
Asarum europaeum	ab	1,772	0.683	1.81	0.843	45.3	0.773	1.82	0.939
Asarum europaeum	ad	2,573	0.652	1.90	0.833	45.3	0.773	1.82	0.939
Peperomia sp.	ad	1,388	0.496	1.86	0.892	13.6	0.304	2.85	0.940

species	side	cell area (µm^2)	aspect ratio	circularity	solidity	leaf area (mm^2)	aspect ratio	circularity	solidity
Drimys winteri	ad	2,378	0.755	1.40	0.933	15.8	0.174	3.82	0.930
Chloranthus sp.	ad	3,030	0.693	2.12	0.805	28.6	0.372	1.97	0.975
Sabal minor	ad	532	0.681	1.37	0.953				
Carex morrowii	ad	1,231	0.274	3.98	0.757				
Avena strigosa	ab	3,541	0.080	7.30	0.846				
Danthonia californica	ab	4,957	0.099	5.59	0.889				
Eriachne sp.	ab	4,278	0.181	4.33	0.795				
Hordeum vulgare	ab	5,499	0.175	3.96	0.887				
Oryza sativa	ad	2,374	0.373	3.69	0.738				
Phalaris arundinacea	NA	988	0.405	2.08	0.904				
Setaria italica	ab	1,180	0.317	2.44	0.900				
Sorghum bicolor	ab	3,205	0.221	4.01	0.821				
Triraphis sp.	ab	6,429	0.172	5.23	0.837				
Triticum aestivum	NA	2,210	0.147	3.97	0.889				
Palisota mannii	ab	5,128	0.642	1.50	0.930	382.8	0.308	2.51	0.968
Palisota mannii	ad	2,660	0.387	1.85	0.939	382.8	0.308	2.51	0.968
Pollia japonica	ab	3,526	0.701	1.43	0.928	31.1	0.337	2.08	0.964
Pollia japonica	ad	2,613	0.594	1.50	0.936	31.1	0.337	2.08	0.964
Tradescantia spathacea sw. 'Vittata'	ab	5,507	0.695	1.39	0.954	74.8	0.175	3.21	0.982
Tradescantia spathacea sw. 'Vittata'	ad	5,433	0.613	1.40	0.960	74.8	0.175	3.21	0.982
Tradescantia virginiana	ab	4,865	0.568	1.89	0.825	22.5	0.085	6.75	0.974
Tradescantia virginiana	ad	5,638	0.611	1.69	0.868	22.5	0.085	6.75	0.974
Eichhornia crassipes	ad	885	0.620	1.54	0.922				
Costus igneus	ab	3,610	0.672	1.43	0.936	42.0	0.533	1.44	0.984
Costus igneus	ad	3,106	0.584	1.51	0.924	42.0	0.533	1.44	0.984
Monocostus uniflorus	ab	2,508	0.766	1.39	0.934	26.2	0.490	1.46	0.989
Monocostus uniflorus	ad	2,365	0.774	1.34	0.950	26.2	0.490	1.46	0.989
Alpinia zerumbet	ad	1,456	0.455	1.75	0.943	18.2	0.450	1.58	0.968
Curcuma longa	ab	1,137	0.588	1.47	0.947	176.2	0.305	2.27	0.987

species	side	cell area (µm^2)	aspect ratio	circularity	solidity	leaf area (mm^2)	aspect ratio	circularity	solidity
Curcuma longa	ad	1,545	0.377	1.70	0.950	176.2	0.305	2.27	0.987
Zingiber officinale Roscoe	ab	674	0.754	1.40	0.929	34.0	0.094	7.80	0.871
Zingiber officinale Roscoe	ad	1,237	0.835	1.33	0.945	34.0	0.094	7.80	0.871
Cirrhopetalum gracillimum	ab	891	0.750	1.36	0.948	18.8	0.228	2.58	0.968
Cirrhopetalum gracillimum	ad	1,030	0.789	1.31	0.953	18.8	0.228	2.58	0.968
Cymbidium sp.	ad	1,584	0.241	2.54	0.898	7.2	0.325	1.81	0.984
Oncidium sphacelatum	ab	637	0.404	1.86	0.916	140.0	0.070	7.53	0.965
Oncidium sphacelatum	ad	870	0.768	1.40	0.943	140.0	0.070	7.53	0.965
Vanilla planifolia	ab	753	0.691	1.35	0.954	64.6	0.355	1.84	0.990
Crocus sativus	ab	3,521	0.131	4.29	0.867				
Crocus sativus	ad	4,395	0.123	4.78	0.846				
Iris germanica 'Kharput'	NA	4,824	0.237	2.91	0.905	92.7	0.063	10.28	0.849
Iris caucasica 'Rainbow Grand'	ab	25,792	0.061	9.61	0.836	11.7	0.075	20.14	0.364
Iris caucasica 'Rainbow Grand'	ad	4,293	0.238	2.91	0.886	11.7	0.075	20.14	0.364
Aloe Vera	ab	5,315	0.626	1.45	0.942				
Haworthia retusa	ab	3,241	0.705	1.36	0.963				
Hemerocallis fulva	ab	3,108	0.174	3.56	0.917	129.1	0.087	15.78	0.446
Hemerocallis fulva	ad	4,916	0.349	2.05	0.904	129.1	0.087	15.78	0.446
Kniphofia caulescens	ab	6,195	0.160	4.00	0.840	120.2	0.172	11.12	0.319
Kniphofia caulescens	ad	8,368	0.123	5.11	0.840	120.2	0.172	11.12	0.319
Agapanthus preacox	ab	6,306	0.210	3.33	0.850				
Amaryllis belladonna	ab	2,542	0.288	2.63	0.859	10.8	0.190	2.78	0.956
Amaryllis belladonna	ad	2,705	0.280	2.54	0.856	10.8	0.190	2.78	0.956
Clivia miniata	ad	912	0.204	3.25	0.887	14.2	0.259	2.36	0.975
Leucojum aestivum	ab	753	0.586	1.69	0.878	6.4	0.109	4.75	0.944
Leucojum aestivum	ad	1,749	0.787	1.36	0.939	6.4	0.109	4.75	0.944
Sternbergia lutea	ab	5,450	0.137	4.80	0.878	9.4	0.116	10.42	0.455
Sternbergia lutea	ad	6,439	0.147	3.97	0.889	9.4	0.116	10.42	0.455
Agave sisalana	ad	1,105	0.425	1.78	0.936				

species	side	cell area (µm^2)	aspect ratio	circularity	solidity	leaf area (mm^2)	aspect ratio	circularity	solidity
Calibanus hookeri	ad	2,885	0.339	1.88	0.959				
Danae racemosa	ab	939	0.444	1.92	0.868	7.9	0.230	2.79	0.954
Danae racemosa	ad	1,816	0.330	2.10	0.916	7.9	0.230	2.79	0.954
Ornithogalum caudatum	ab	26,967	0.093	6.45	0.879				
Rohdea japonica	ab	3,148	0.545	1.63	0.910	90.4	0.350	1.98	0.991
Rohdea japonica	ad	3,565	0.585	1.54	0.929	90.4	0.350	1.98	0.991
Ruscus aculeatus	ab	1,056	0.565	1.64	0.910	0.8	0.565	1.57	0.950
Ruscus aculeatus	ad	971	0.501	1.88	0.870	0.8	0.565	1.57	0.950
Ruscus hypoglossum	ad	1,935	0.467	2.13	0.845	0.1	0.146	4.23	0.873
Yucca gloriosa	ab	575	0.679	1.43	0.949	86.9	0.130	4.76	0.945
Yucca gloriosa	ad	620	0.696	1.40	0.953	86.9	0.130	4.76	0.945
Alstroemeria aurea	ab	14,240	0.609	4.08	0.709	6.1	0.191	3.33	0.948
Alstroemeria aurea	ad	16,575	0.274	3.17	0.807	6.1	0.191	3.33	0.948
Luzuriaga latifolia	ab	1,417	0.683	1.42	0.944	14.3	0.091	5.77	0.975
Smilax bona-nox 'Cantab'	ab	1,009	0.594	2.43	0.783	34.3	0.803	1.94	0.932
Smilax bona-nox 'Cantab'	ad	2,268	0.688	3.15	0.744	34.3	0.803	1.94	0.932
Tricyrtis hirta	ab	3,222	0.677	2.41	0.763	21.9	0.350	2.31	0.941
Tricyrtis hirta	ad	4,587	0.671	1.57	0.895	21.9	0.350	2.31	0.941
Dioscorea bulbifera	ad	2,610	0.613	1.96	0.828	19.6	0.294	2.20	0.980
Carludovica palmata	ab	1,357	0.730	1.45	0.918				
Freycinetia sp.	ad	2,432	0.801	1.32	0.953				
Pandanus veitchei	ab	845	0.147	3.91	0.905				
Stichoneuron caudatum	ad	1,417	0.683	1.42	0.943	0.3	0.074	5.74	0.939
Aglaonema crispum	ab	3,655	0.608	1.72	0.867	299.2	0.495	1.36	0.997
Aglaonema crispum	ad	3,254	0.721	1.99	0.826	299.2	0.495	1.36	0.997
Amorphaphallus titanum	ab	4,075	0.345	1.92	0.936				
Arum maculatum	ab	2,001	0.608	1.80	0.841	56.2	0.539	1.79	0.930
Arum maculatum	ad	2,802	0.650	1.60	0.884	56.2	0.539	1.79	0.930
Dracunculus canariensis	ab	2,279	0.583	1.78	0.861	18.9	0.504	1.44	0.990

species	side	cell area (µm^2)	aspect ratio	circularity	solidity	leaf area (mm^2)	aspect ratio	circularity	solidity
Dracunculus canariensis	ad	3,194	0.633	1.64	0.892	18.9	0.504	1.44	0.990
Rhektophyllum mirabile	ab	1,489	0.624	1.63	0.881	298.5	0.814	2.30	0.835
Rhektophyllum mirabile	ad	1,126	0.589	1.76	0.857	298.5	0.814	2.30	0.835
Spathiphyllum wallisii	ab	1,394	0.606	2.25	0.782	118.9	0.314	2.05	0.978
Spathiphyllum wallisii	ad	1,443	0.699	2.47	0.779	118.9	0.314	2.05	0.978
Zantedeschia aethiopica	ab	1,138	0.572	1.78	0.858	192.8	0.597	2.11	0.897
Zantedeschia aethiopica	ad	1,253	0.674	1.49	0.914	192.8	0.597	2.11	0.897
Hydrocleys nymphoides	ab	753	0.593	1.60	0.904				
Acorus gramineus	ab	382	0.214	3.25	0.869				
Aconitum carmichaelii var. arendsi	ab	3,456	0.660	2.67	0.766	19.6	0.534	3.10	0.853
Aconitum carmichaelii var. arendsi	ad	5,667	0.693	1.91	0.842	19.6	0.534	3.10	0.853
Anemone canadensis	ab	1,905	0.645	2.78	0.747	11.6	0.477	3.48	0.860
Anemone canadensis	ad	2,112	0.656	1.61	0.895	11.6	0.477	3.48	0.860
Helleborus orientalis ssp orientalis	ab	2,917	0.708	2.21	0.807	20.9	0.369	2.62	0.956
Helleborus orientalis ssp orientalis	ad	4,587	0.731	2.26	0.817	20.9	0.369	2.62	0.956
Astragalus falcatus	ab	1,572	0.701	1.97	0.816	1.9	0.345	1.88	0.983
Astragalus falcatus	ad	3,226	0.693	1.77	0.865	1.9	0.345	1.88	0.983
Dorycinum rectum	ab	913	0.725	1.89	0.850	4.0	0.447	1.74	0.983
Dorycinum rectum	ad	777	0.692	1.43	0.931	4.0	0.447	1.74	0.983
Medicago sativa	ab	1,277	0.584	2.20	0.826	0.7	0.417	1.79	0.966
Medicago sativa	ad	913	0.594	1.92	0.852	0.7	0.417	1.79	0.966
Mimosa pudica	ad	341	0.602	1.86	0.820	49.3	0.332	2.14	0.976
Senna didymobotrya	ad	1,906	0.649	3.11	0.674	38.9	0.198	7.14	0.592
Trifolium pannonicum	ab	2,450	0.676	1.69	0.886	1.8	0.506	1.53	0.981
Trifolium pannonicum	ad	2,228	0.636	1.54	0.923	1.8	0.506	1.53	0.981
Ulex europaeus	NA	360	0.433	2.33	0.817	0.0	0.112	5.51	0.842
Geum triflorum 'Purple Avens'	ab	662	0.579	2.28	0.791	69.5	0.964	3.29	0.894
Geum triflorum 'Purple Avens'	ad	592	0.684	1.51	0.912	69.5	0.964	3.29	0.894
Rosa damascena 'versicolor'	ab	394	0.683	1.65	0.866	8.5	0.664	2.05	0.960

species	side	cell area (µm^2)	aspect ratio	circularity	solidity	leaf area (mm^2)	aspect ratio	circularity	solidity
Rosa damascena 'versicolor'	ad	632	0.712	1.42	0.936	8.5	0.664	2.05	0.960
Lithocarpus henryi	ad	2,052	0.175	3.89	0.855				
Oxalis regnellii	ad	8,158	0.727	1.71	0.867	265.5	0.539	5.24	0.801
Oxalis tetraphylla	NA	1,772	0.798	1.35	0.949	6.7	0.724	1.48	0.976
Oxalis valdiviensis	ab	1,119	0.747	1.49	0.911	0.8	0.969	1.44	0.943
Oxalis valdiviensis	ad	986	0.724	1.65	0.877	0.8	0.969	1.44	0.943
Hiptage benghalensis	ad	789	0.604	1.89	0.853	3.8	0.529	1.41	0.980
Passiflora edulis	ad	2,579	0.630	1.79	0.854	25.9	0.075	5.92	0.984
Passiflora sp.	ad	1,849	0.685	2.15	0.797	9.5	0.493	1.56	0.962
Poliothyrsis sinensis	NA	1,173	0.651	2.36	0.784	54.7	0.424	1.89	0.936
Euphorbia flanaganii	ad	2,261	0.684	1.68	0.882	34.4	0.541	1.85	0.964
Euphorbia mellifera	ab	401	0.806	1.45	0.910	36.9	0.178	3.21	0.973
Euphorbia mellifera	ad	970	0.749	1.52	0.901	36.9	0.178	3.21	0.973
Euphorbia pulcherrima	ad	3,079	0.715	3.24	0.697	4.0	0.328	2.52	0.890
Catha edulis	ad	944	0.644	3.22	0.688	11.3	0.576	1.62	0.965
Guaiacum officinale	ad	444	0.640	1.40	0.957	22.6	0.600	1.42	0.982
Pelargonium carnosum	ad	2,679	0.688	1.50	0.915	1.5	0.910	1.26	0.965
Cuphea ignea	ab	1,096	0.752	3.50	0.682	4.6	0.353	1.93	0.982
Cuphea ignea	ad	1,427	0.684	1.81	0.839	4.6	0.353	1.93	0.982
Heimia myrthifolia	ab	436	0.679	2.03	0.800	1.7	0.221	3.13	0.963
Heimia myrthifolia	ad	774	0.668	1.45	0.932	1.7	0.221	3.13	0.963
Fuchsia magellanica var. molinae	ab	1,005	0.686	4.28	0.644	6.5	0.449	1.74	0.955
Fuchsia magellanica var. molinae	ad	1,343	0.617	1.73	0.868	6.5	0.449	1.74	0.955
Fuchsia 'Mrs Popple'	NA	1,365	0.717	3.47	0.703	10.4	0.653	1.50	0.965
Oenothera stricta	ab	1,810	0.657	3.16	0.718	4.4	0.201	4.03	0.858
Oenothera stricta	ad	1,994	0.672	2.71	0.759	4.4	0.201	4.03	0.858
Ceiba pentandra	ad	2,558	0.703	1.57	0.912	78.9	0.486	1.46	0.994
Malva sylvestris	ab	382	0.662	2.62	0.688	3.8	0.944	1.82	0.907
Malva sylvestris	ad	486	0.673	1.96	0.815	3.8	0.944	1.82	0.907

species	side	cell area (µm^2)	aspect ratio	circularity	solidity	leaf area (mm^2)	aspect ratio	circularity	solidity
Carica papaya	ad	2,550	0.704	1.59	0.894				
Erysimum scoparium	ab	831	0.666	2.06	0.796	2.7	0.123	6.56	0.661
Erysimum scoparium	ad	1,272	0.617	1.66	0.885	2.7	0.123	6.56	0.661
Sisymbrium austriaticum	ab	1,025	0.686	1.91	0.803	2.4	0.509	1.80	0.943
Sisymbrium austriaticum	ad	756	0.632	2.34	0.752	2.4	0.509	1.80	0.943
Bursera schlechtendalii	ad	730	0.694	1.52	0.914	25.0	0.533	1.82	0.967
Rhus potaninii	ab	702	0.682	1.46	0.927	15.0	0.395	1.88	0.981
Rhus potaninii	ad	2,165	0.718	1.46	0.907	15.0	0.395	1.88	0.981
Dodonaea viscosa	ad	784	0.547	1.86	0.861	0.2	0.087	5.09	0.943
Citrus x limon	ab	286	0.675	2.40	0.765	6.6	0.213	2.88	0.954
Citrus x limon	ad	770	0.727	2.33	0.782	6.6	0.213	2.88	0.954
Murraya koenigii	ad	673	0.574	1.60	0.915	4.5	0.198	3.03	0.932
Cissus quadrangularis	ab	1,533	0.571	1.58	0.941				
Cissus tuberosa	ad	2,126	0.619	1.62	0.896	0.0	0.166	3.08	0.928
Paeonia tenuifolia	ab	2,445	0.638	3.42	0.725	8.5	0.221	3.21	0.941
Paeonia tenuifolia	ad	3,248	0.730	2.79	0.740	8.5	0.221	3.21	0.941
Bergenia purpurascens	ab	1,116	0.564	1.74	0.877	110.4	0.578	1.47	0.989
Bergenia purpurascens	ad	1,487	0.723	1.37	0.948	110.4	0.578	1.47	0.989
Saxifraga canaliculata	ab	1,292	0.466	1.89	0.883	0.4	0.497	5.12	0.612
Saxifraga canaliculata	ad	1,480	0.541	1.83	0.858	0.4	0.497	5.12	0.612
Saxifraga hostii ssp hostii	ab	1,307	0.496	2.00	0.844	3.1	0.153	4.78	0.833
Saxifraga hostii ssp hostii	ad	2,175	0.726	1.41	0.942	3.1	0.153	4.78	0.833
Adromischus sp.	ad	6,139	0.570	1.59	0.919	2.8	0.848	77.45	0.134
Crassula fasciatudform	ab	3,894	0.530	1.62	0.935				
Haloragis erecta	ab	447	0.468	2.30	0.796	4.7	0.494	3.52	0.883
Haloragis erecta	ad	724	0.703	1.46	0.928	4.7	0.494	3.52	0.883
Ceratostigma plumbaginoides	ab	872	0.626	1.86	0.834	13.3	0.600	2.63	0.901
Ceratostigma plumbaginoides	ad	755	0.663	1.50	0.908	13.3	0.600	2.63	0.901
Ceratostigma wilmottianum	ab	1,280	0.588	2.09	0.810	5.9	0.459	2.52	0.868

species	side	cell area (µm^2)	aspect ratio	circularity	solidity	leaf area (mm^2)	aspect ratio	circularity	solidity
Ceratostigma wilmottianum	ad	1,459	0.678	1.50	0.912	5.9	0.459	2.52	0.868
Homalocladium platycladum	ad	502	0.629	3.33	0.670	21.1	0.390	1.72	0.985
Oxyria digyna	ab	2,372	0.588	3.04	0.704	14.9	0.953	1.53	0.951
Oxyria digyna	ad	2,759	0.593	1.94	0.831	14.9	0.953	1.53	0.951
Persicaria affinis	ab	1,293	0.613	1.81	0.842	6.5	0.159	4.06	0.921
Persicaria affinis	ad	1,282	0.661	1.45	0.935	6.5	0.159	4.06	0.921
Persicaria polystachya	ab	626	0.655	3.76	0.669	119.4	0.379	2.45	0.934
Persicaria polystachya	ad	1,561	0.613	1.76	0.866	119.4	0.379	2.45	0.934
Persicaria weyrichii	ab	2,383	0.682	1.41	0.942	58.7	0.261	3.27	0.892
Persicaria weyrichii	ad	1,035	0.619	2.06	0.802	58.7	0.261	3.27	0.892
Rumex acetosella	ab	449	0.572	2.33	0.745	11.4	0.384	2.20	0.921
Rumex acetosella	ad	880	0.732	1.53	0.898	11.4	0.384	2.20	0.921
Rumex scutatus	ab	1,063	0.656	3.14	0.671	16.7	0.954	2.31	0.806
Rumex scutatus	ad	1,492	0.682	2.13	0.778	16.7	0.954	2.31	0.806
Amaranthus caudatus	ab	1,623	0.620	4.74	0.638	36.2	0.610	1.42	0.982
Amaranthus caudatus	ad	1,266	0.625	1.76	0.857	36.2	0.610	1.42	0.982
Amaranthus hybridus	ab	1,500	0.640	3.15	0.702	34.1	0.524	1.59	0.981
Amaranthus hybridus	ad	1,537	0.648	1.79	0.858	34.1	0.524	1.59	0.981
Beta trigyna	ab	1,271	0.648	2.35	0.774	68.5	0.720	1.63	0.963
Beta trigyna	ad	1,545	0.729	1.63	0.880	68.5	0.720	1.63	0.963
Beta vulgaris ssp. vulgaris	ab	741	0.663	1.76	0.850	42.1	0.548	1.73	0.952
Beta vulgaris ssp. vulgaris	ad	986	0.668	1.65	0.871	42.1	0.548	1.73	0.952
Chenopodium bonus-henricus	ab	999	0.611	3.05	0.709	46.9	0.719	2.14	0.843
Chenopodium bonus-henricus	ad	1,072	0.590	1.78	0.873	46.9	0.719	2.14	0.843
Ercilla volubilis	ab	437	0.545	1.90	0.828	6.8	0.547	1.64	0.959
Ercilla volubilis	ad	916	0.724	1.43	0.934	6.8	0.547	1.64	0.959
Epiphyllum sp.	ab	1,393	0.775	1.39	0.958				
Impatiens repens	ad	4,368	0.711	3.73	0.681	54.7	0.434	1.61	0.990
Ardisia crispa	ad	1,898	0.665	3.33	0.688	2.1	0.340	2.03	0.973

species	side	cell area (µm^2)	aspect ratio	circularity	solidity	leaf area (mm^2)	aspect ratio	circularity	solidity
Macleania insignis	ad	2,713	0.779	3.74	0.637	38.2	0.412	1.66	0.983
Ilex paraguariensis	ad	1,160	0.546	3.61	0.660	86.3	0.706	1.79	0.925
Campanula fenestrellata	ab	774	0.654	2.01	0.799	10.7	1.015	2.44	0.858
Campanula fenestrellata	ad	1,338	0.718	1.52	0.903	10.7	1.015	2.44	0.858
Campanula poscharskyana	ab	718	0.674	2.15	0.790	6.5	0.654	3.25	0.904
Campanula poscharskyana	ad	2,279	0.735	1.42	0.928	6.5	0.654	3.25	0.904
Jasione heldreichii	ab	2,148	0.763	3.18	0.721	1.4	0.123	5.60	0.862
Jasione heldreichii	ad	3,379	0.735	2.99	0.738	1.4	0.123	5.60	0.862
Berkheya purpurea	ab	1,033	0.692	2.57	0.772	24.1	0.243	23.39	0.651
Berkheya purpurea	ad	2,007	0.635	2.00	0.832	24.1	0.243	23.39	0.651
Berkheya radula	NA	1,452	0.666	1.75	0.868	4.4	0.515	1.61	0.943
Centaurea dealbata	ab	733	0.653	1.96	0.818	1.6	0.436	3.37	0.703
Centaurea dealbata	ad	659	0.626	1.67	0.881	1.6	0.436	3.37	0.703
Centaurea simplicicaulis	NA	670	0.702	1.56	0.904	1.6	0.779	1.90	0.910
Othonna crassifolia	ad	2,150	0.719	1.68	0.868	12.4	0.866	2.04	0.866
Picris echioides	ab	880	0.608	1.83	0.850	18.8	0.217	3.33	0.937
Picris echioides	ad	705	0.701	1.51	0.907	18.8	0.217	3.33	0.937
Scorzonera hispanica	ab	1,251	0.688	1.81	0.843	32.1	0.301	3.12	0.871
Scorzonera hispanica	ad	1,224	0.620	1.71	0.874	32.1	0.301	3.12	0.871
Hedera nepalensis	ad	1,140	0.587	2.14	0.793	108.4	0.139	3.88	0.982
Bupleurum fruticosum	ab	681	0.692	1.66	0.869	10.6	0.263	2.22	0.989
Bupleurum fruticosum	ad	894	0.763	1.38	0.934	10.6	0.263	2.22	0.989
Eryngium agavifolium	ab	546	0.618	1.71	0.876	33.9	0.183	27.61	0.720
Eryngium agavifolium	ad	544	0.673	1.52	0.907	33.9	0.183	27.61	0.720
Eryngium bourgatii	ab	490	0.652	1.83	0.810	2.7	0.710	8.53	0.460
Eryngium bourgatii	ad	547	0.743	1.62	0.863	2.7	0.710	8.53	0.460
Myrrhis odorata	ab	704	0.663	3.09	0.698	3.7	0.645	16.79	0.754
Myrrhis odorata	ad	1,380	0.658	2.30	0.771	3.7	0.645	16.79	0.754
Pastinaca sativa 'Tender and true'	ab	949	0.674	4.01	0.630	41.2	0.729	3.40	0.837

species	side	cell area (µm^2)	aspect ratio	circularity	solidity	leaf area (mm^2)	aspect ratio	circularity	solidity
Pastinaca sativa 'Tender and true'	ad	2,012	0.683	1.88	0.839	41.2	0.729	3.40	0.837
Centrathus ruber	ab	927	0.687	2.35	0.778	3.7	0.202	3.61	0.906
Centrathus ruber	ad	1,007	0.603	1.65	0.889	3.7	0.202	3.61	0.906
Cephalaria flava	ab	1,818	0.652	3.34	0.686	1.8	0.451	2.40	0.918
Cephalaria flava	ad	2,355	0.718	1.84	0.838	1.8	0.451	2.40	0.918
Lonicera quinquelocularis	NA	852	0.720	1.47	0.920	36.3	0.622	1.44	0.986
Scabiosa olgae	ab	1,774	0.721	1.46	0.920	10.8	0.517	1.64	0.955
Scabiosa olgae	ad	1,801	0.693	1.48	0.922	10.8	0.517	1.64	0.955
Succisella inflexa	ab	389	0.682	1.94	0.794	25.1	0.185	3.60	0.953
Succisella inflexa	ad	547	0.669	1.71	0.857	25.1	0.185	3.60	0.953
Valeriana phu	NA	2,371	0.667	3.13	0.718	25.6	0.433	1.68	0.981
Brugmansia suaveolens	ad	2,261	0.728	1.89	0.839	10.4	0.320	1.83	0.984
Capsicum annuum 'Etna'	ab	638	0.639	2.77	0.696	6.0	0.532	1.56	0.964
Capsicum annuum 'Etna'	ad	1,025	0.629	2.00	0.801	6.0	0.532	1.56	0.964
Capsicum annuum 'Mohawk'	ab	2,063	0.644	3.08	0.711	6.5	0.710	1.34	0.988
Capsicum annuum 'Mohawk'	ad	5,065	0.645	1.80	0.857	6.5	0.710	1.34	0.988
Salpichroa origanifolia	ab	5,595	0.627	2.32	0.779	12.2	0.776	1.32	0.988
Salpichroa origanifolia	ad	6,106	0.665	2.28	0.782	12.2	0.776	1.32	0.988
Montinia caryophyllacea	ad	1,023	0.661	1.68	0.885	19.5	0.276	2.37	0.971
Jasminum fruticans	NA	1,209	0.799	2.04	0.835	2.2	0.444	1.78	0.950
Jasminum humile var. revolutum	ab	258	0.708	1.94	0.814	4.9	0.397	1.73	0.978
Jasminum humile var. revolutum	ad	570	0.718	1.87	0.846	4.9	0.397	1.73	0.978
Syringa 'Charles X'	ab	912	0.609	1.84	0.857	51.5	0.834	1.36	0.961
Syringa 'Charles X'	ad	2,055	0.749	1.46	0.911	51.5	0.834	1.36	0.961
Syringa vulgaris	ab	538	0.643	1.66	0.873	80.5	0.794	1.34	0.966
Syringa vulgaris	ad	1,153	0.704	1.43	0.939	80.5	0.794	1.34	0.966
Globularia punctata	NA	948	0.705	2.87	0.732	4.7	0.640	1.44	0.975
Globularia trichosantha	ab	1,099	0.683	2.52	0.764	2.8	0.510	1.60	0.973
Globularia trichosantha	ad	837	0.691	2.94	0.714	2.8	0.510	1.60	0.973

species	side	cell area (µm^2)	aspect ratio	circularity	solidity	leaf area (mm^2)	aspect ratio	circularity	solidity
Hebe rakaiensis	ab	529	0.680	1.58	0.888	1.0	0.358	1.88	0.978
Hebe rakaiensis	ad	933	0.728	1.45	0.912	1.0	0.358	1.88	0.978
Penstemon 'Raven'	ab	636	0.707	2.54	0.751	22.6	0.308	2.24	0.963
Penstemon 'Raven'	ad	786	0.760	1.55	0.895	22.6	0.308	2.24	0.963
Plantago cynops	ab	1,179	0.674	2.17	0.793	0.9	0.091	37.41	0.211
Plantago cynops	ad	1,110	0.606	1.80	0.876	0.9	0.091	37.41	0.211
Synthyris missurica var stellata	ab	1,081	0.714	2.12	0.810	29.0	0.922	2.16	0.926
Synthyris missurica var stellata	ad	2,222	0.756	2.01	0.828	29.0	0.922	2.16	0.926
Veronica gentianoides	NA	2,638	0.693	2.10	0.811	15.6	0.365	1.96	0.983
Veronica petraea	ab	1,457	0.696	3.62	0.703	1.8	0.222	2.91	0.946
Veronica petraea	ad	2,484	0.749	2.22	0.814	1.8	0.222	2.91	0.946
Scrophularia canina ssp. Bicolor	NA	1,379	0.711	1.95	0.824	1.7	0.514	2.70	0.845
Scrophularia heterophylla	ab	951	0.680	1.80	0.839	3.2	0.515	2.14	0.911
Scrophularia heterophylla	ad	741	0.605	2.57	0.719	3.2	0.515	2.14	0.911
Acanthus hungaricus	ab	1,156	0.648	2.96	0.664	17.1	0.566	2.82	0.804
Acanthus hungaricus	ad	1,283	0.729	1.68	0.868	17.1	0.566	2.82	0.804
Acanthus spinosus	ab	1,118	0.608	2.72	0.697	6.1	0.420	4.26	0.695
Acanthus spinosus	ad	912	0.635	1.94	0.820	6.1	0.420	4.26	0.695
Justicia guttata	ad	2,636	0.686	1.76	0.860	2.1	0.239	2.45	0.977
Verbena bonariensis	NA	1,514	0.636	2.16	0.788	11.5	0.172	6.28	0.839
Callicarpa japonica	ab	791	0.628	3.47	0.674	12.7	0.383	2.60	0.927
Callicarpa japonica	ad	965	0.675	2.32	0.777	12.7	0.383	2.60	0.927
Clerodendrum thomsoniae	ad	2,156	0.713	4.92	0.612	11.1	0.521	1.48	0.984
Lamium orvala	ab	904	0.626	3.50	0.643	67.6	0.728	3.40	0.882
Lamium orvala	ad	1,985	0.606	2.24	0.797	67.6	0.728	3.40	0.882
Mentha piperita	ab	1,027	0.679	4.13	0.608	1.3	0.494	1.55	0.976
Mentha piperita	ad	1,340	0.736	1.60	0.882	1.3	0.494	1.55	0.976
Salvia officinalis 'Purpurascens'	NA	655	0.655	1.81	0.853	10.1	0.432	2.02	0.975
Stachys macrantha	ab	1,085	0.653	2.87	0.704	9.2	0.722	2.70	0.908

species	side	cell area (µm^2)	aspect ratio	circularity	solidity	leaf area (mm^2)	aspect ratio	circularity	solidity
Stachys macrantha	ad	1,468	0.615	1.79	0.859	9.2	0.722	2.70	0.908
Galium odoratum	ab	3,253	0.641	3.77	0.650	1.3	0.308	2.28	0.931
Galium odoratum	ad	5,121	0.703	2.52	0.772	1.3	0.308	2.28	0.931
Galium rubioides	ab	1,984	0.640	3.10	0.686	2.9	0.214	2.84	0.948
Galium rubioides	ad	2,975	0.767	2.25	0.802	2.9	0.214	2.84	0.948
Rubia tinctorum	ab	1,633	0.598	2.81	0.705	3.3	0.218	3.01	0.933
Rubia tinctorum	ad	1,474	0.593	2.56	0.741	3.3	0.218	3.01	0.933
Fagraea berteriana	ad	2,050	0.599	2.38	0.754	8.8	0.313	2.77	0.867
Catharanthus roseus	ad	1,134	0.722	1.42	0.942	48.9	0.635	1.65	0.884
Ceropegia sandersonii	ad	187	0.667	1.48	0.916				
Rauvolfia verticillata	ad	382	0.633	2.26	0.781	49.0	0.166	3.58	0.956
Tabernaemontana divaricata	ad	1,163	0.740	1.31	0.967	55.6	0.671	1.31	0.985
Anchusa officinalis	ab	1,750	0.695	3.47	0.707	14.0	0.398	2.47	0.925
Anchusa officinalis	ad	2,124	0.690	2.49	0.788	14.0	0.398	2.47	0.925
Borago officinalis	ab	1,046	0.656	2.22	0.786	15.2	0.449	2.09	0.950
Borago officinalis	ad	1,172	0.584	1.84	0.847	15.2	0.449	2.09	0.950
Cerinthe minor	NA	1,348	0.694	1.96	0.822	18.7	0.263	3.27	0.848
Cordia nitida	ad	812	0.589	3.13	0.709	7.5	0.479	1.57	0.980
Nemophila menziesii	NA	4,398	0.736	3.60	0.716	3.5	0.471	3.91	0.689
Symphytum caucasium x S. orientale	ab	744	0.717	3.34	0.686	72.1	0.540	1.66	0.972
Symphytum caucasium x S. orientale	ad	2,017	0.626	1.97	0.831	72.1	0.540	1.66	0.972
Symphytum grandiflorum	ab	1,076	0.669	2.57	0.737	17.3	0.783	1.26	0.987
Symphytum grandiflorum	ad	2,204	0.669	1.45	0.936	17.3	0.783	1.26	0.987

Appendix D

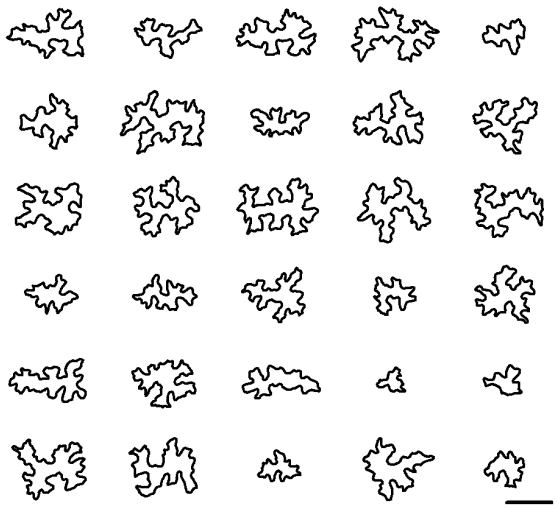
Cell shape data from *Arabidopsis lines*

D.1 Cell outlines

The segmented cell shapes from *Arabidopsis* lines used in the analysis discussed in Chapter 5 are presented on the following pages in Figures D.1–D.21, arranged per cotyledon imaged. Three cotyledons have been imaged for each of the 19 MAGIC parent lines and one cotyledon each for the six mutant lines studied. The scalebar marks 100 µm.

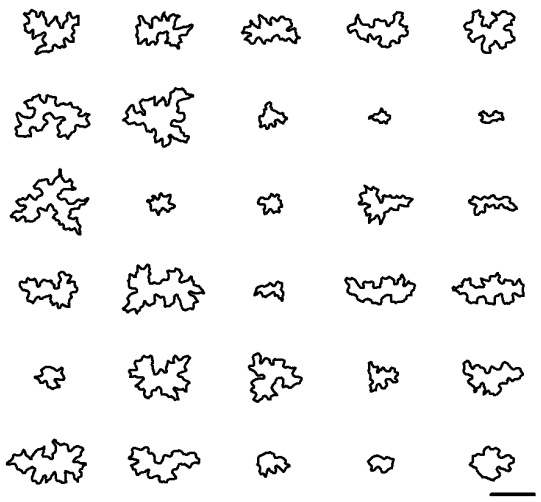
D.2 Cell descriptors

Figures show the distribution of area, aspect ratio, circularity and solidity values for every sample. Median values and quartiles are also marked. Different panels show different MAGIC parent lines, except for the last one that is dedicated to the mutant lines studied. Data in each panel is grouped by the seedlings the cells were taken from. These results are discussed in Section 5.4.



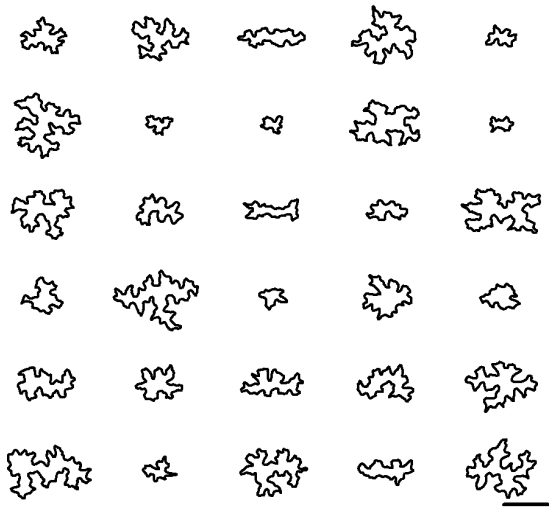
(a) Seedling 1

(b) Seedling 2

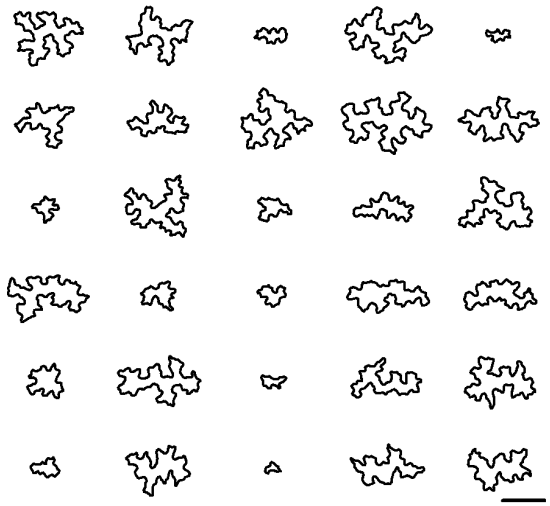


(c) Seedling 3

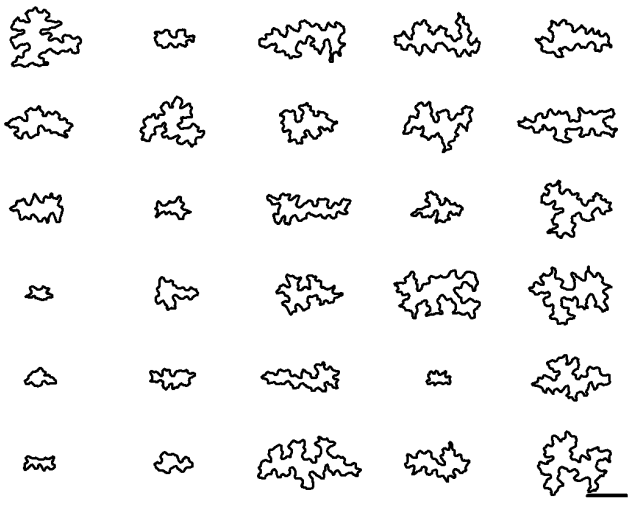
Figure D.1 – Bur-0



(a) Seedling 1

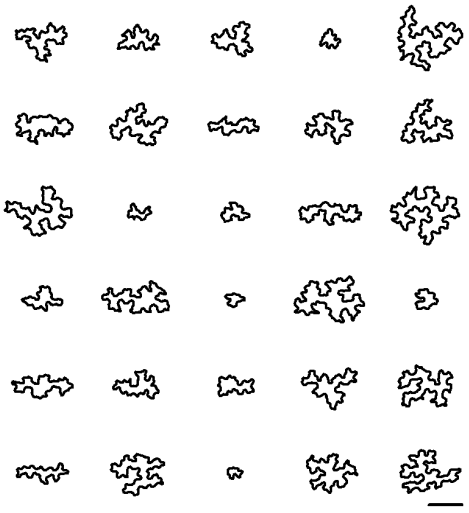


(b) Seedling 2

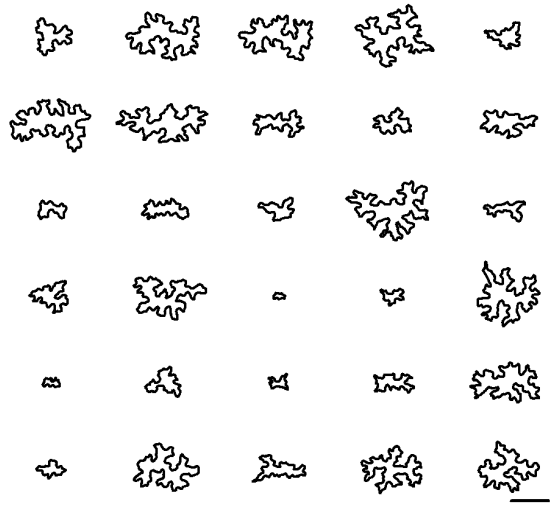


(c) Seedling 3

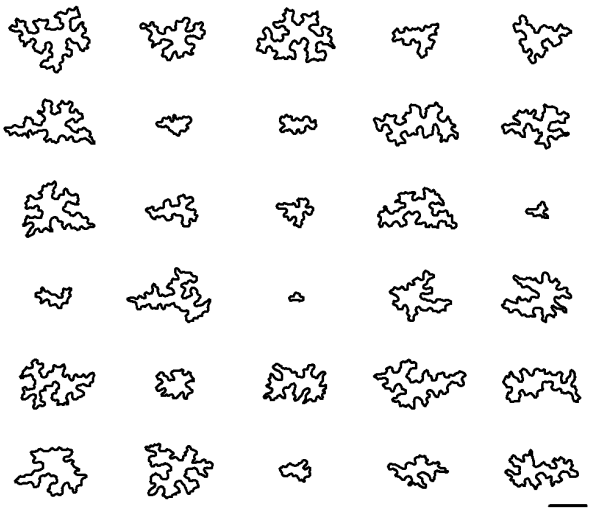
Figure D.2 – Can-0



(a) Seedling 1

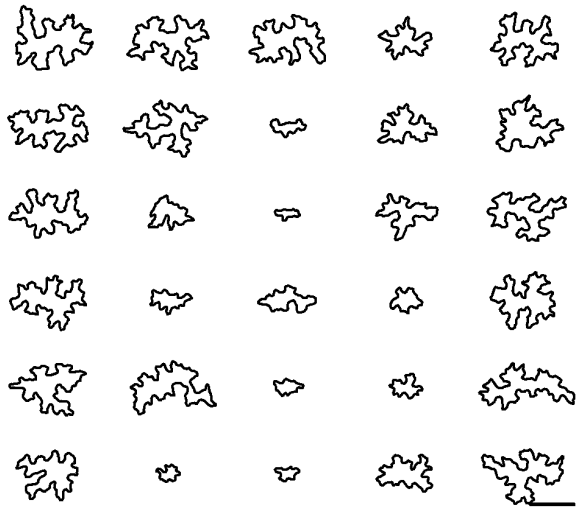


(b) Seedling 2

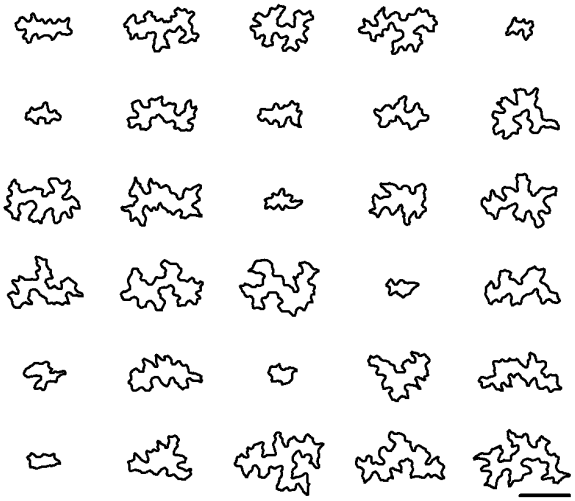


(c) Seedling 3

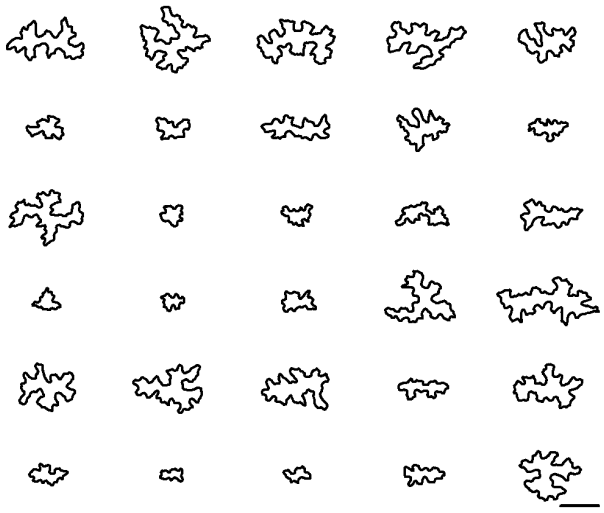
Figure D.3 – Col-0



(a) Seedling 1

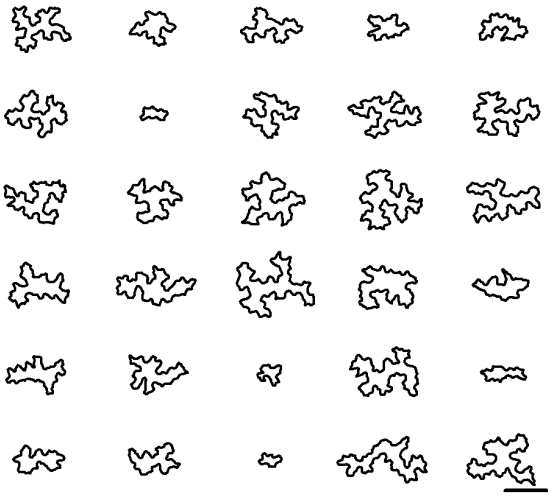


(b) Seedling 2

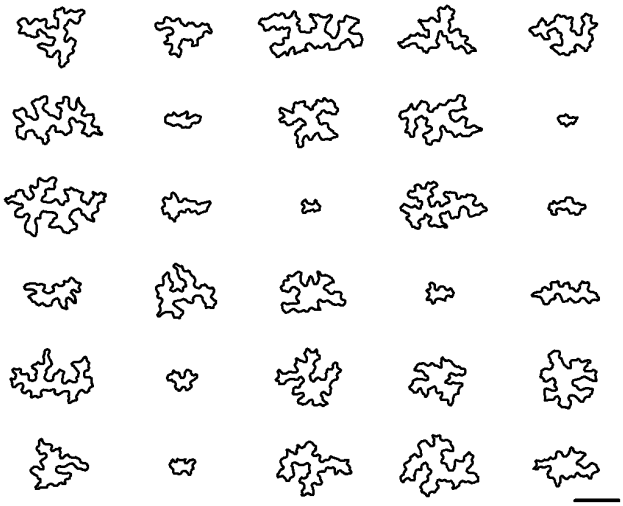


(c) Seedling 3

Figure D.4 – Ct-1



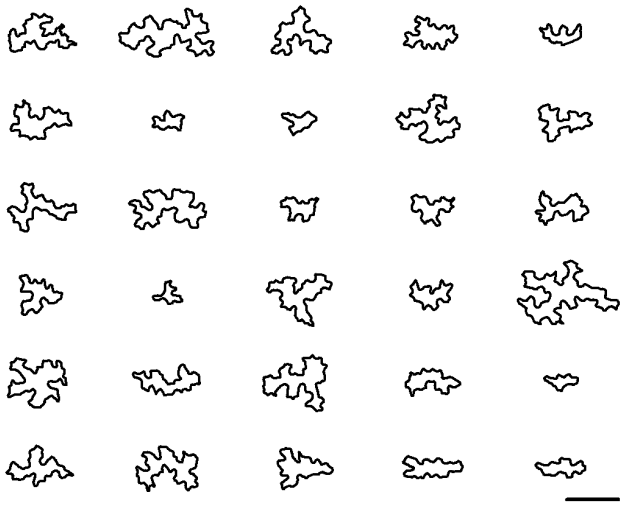
(a) Seedling 1



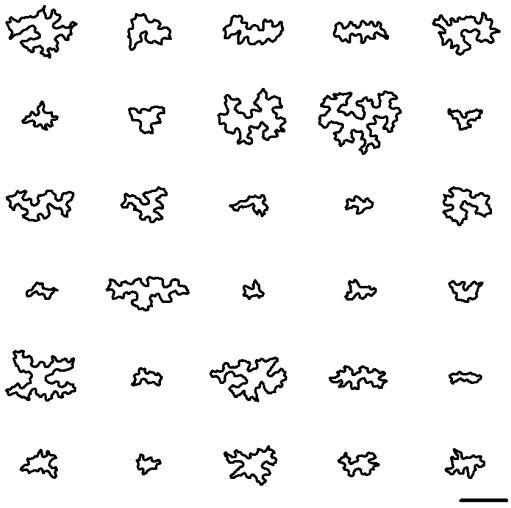
(b) Seedling 2

(c) Seedling 3

Figure D.5 – Edi-1



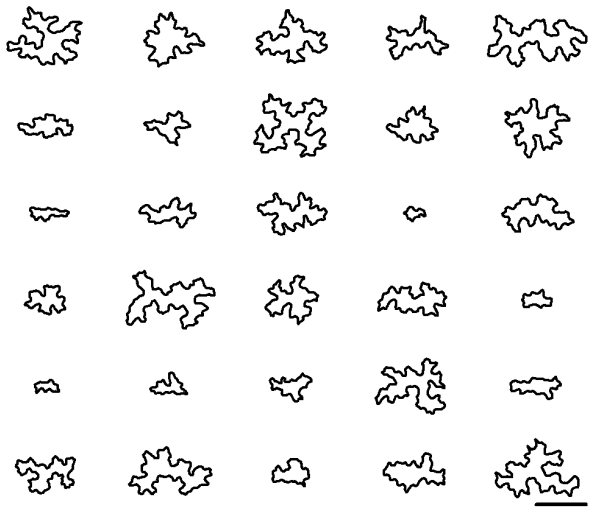
(a) Seedling 1



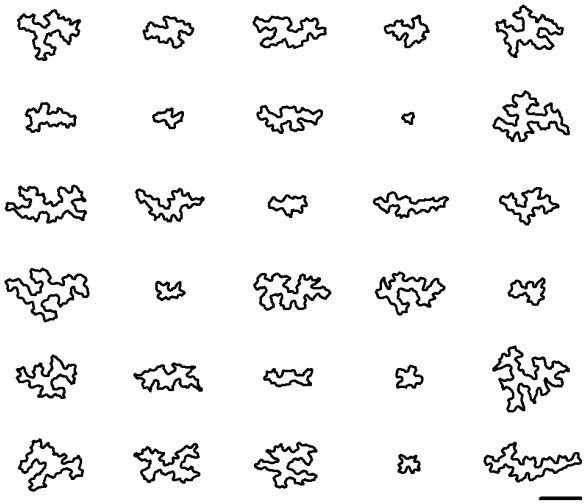
(b) Seedling 2

(c) Seedling 3

Figure D.6 – Hi-0



(a) Seedling 1

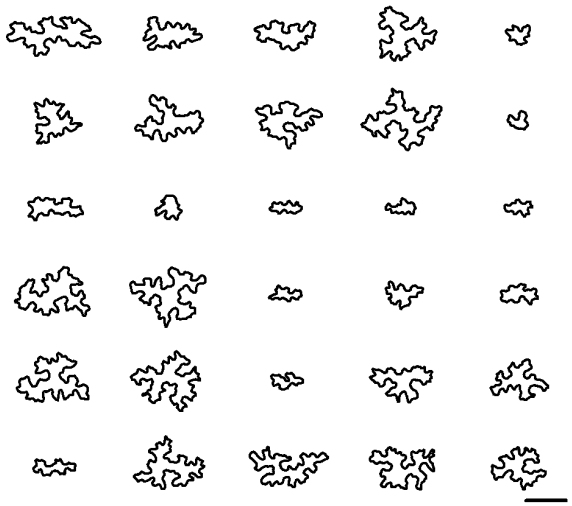


(b) Seedling 2

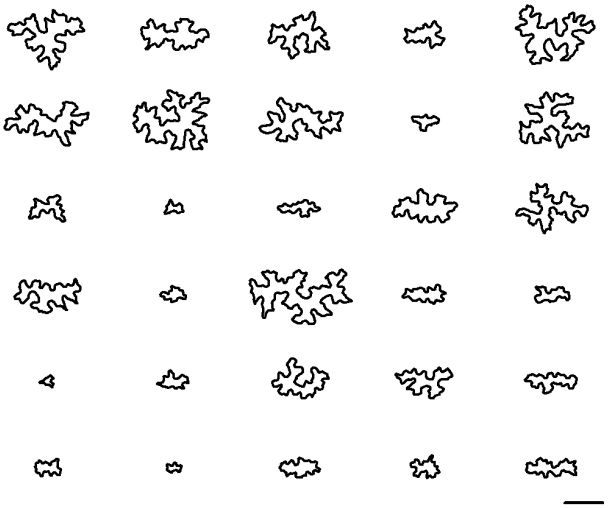
(c) Seedling 3

Figure D.7 – Kn-0

(a) Seedling 1

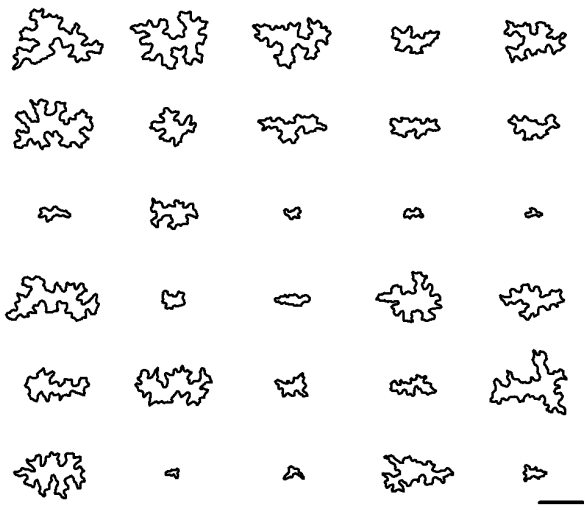


(b) Seedling 2

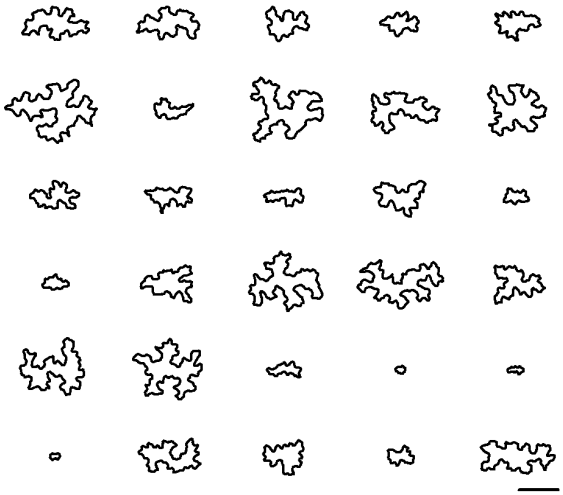


(c) Seedling 3

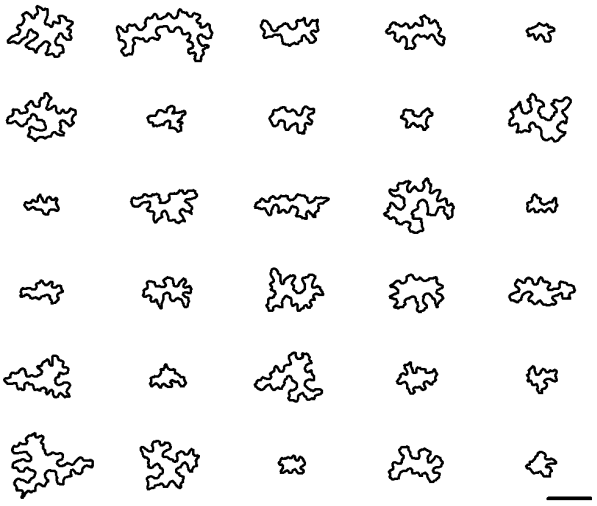
Figure D.8 – Ler-0



(a) Seedling 1

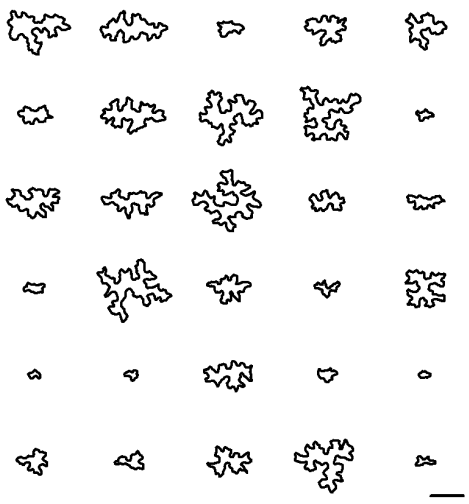


(b) Seedling 2

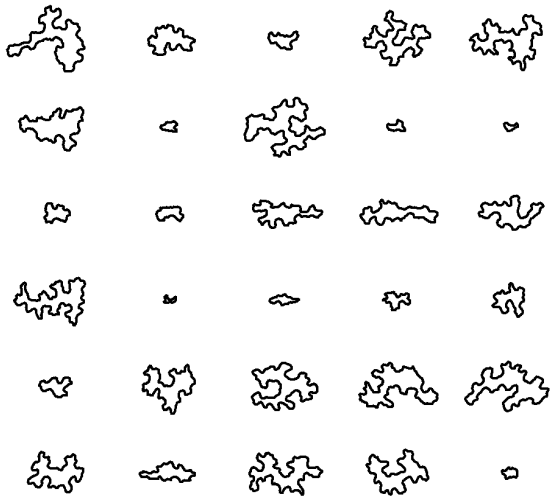


(c) Seedling 3

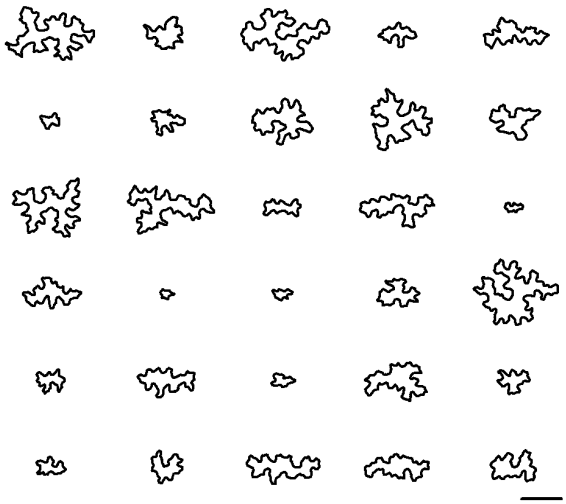
Figure D.9 – Mt-0



(a) Seedling 1



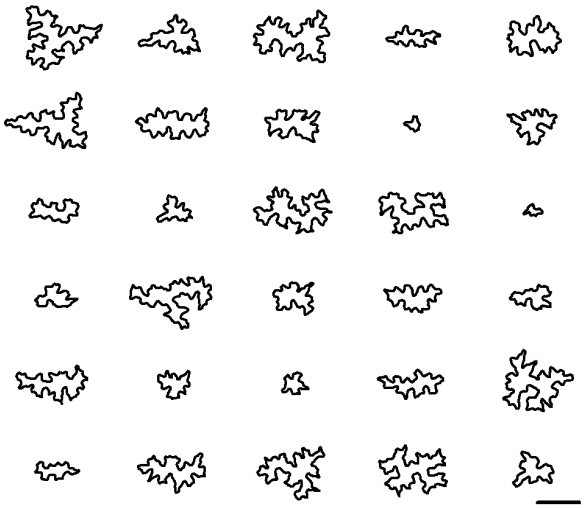
(b) Seedling 2



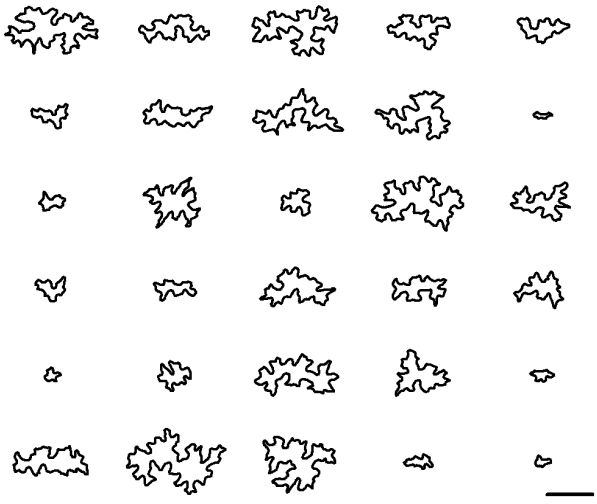
(c) Seedling 3

Figure D.10 – No-0

(a) Seedling 1

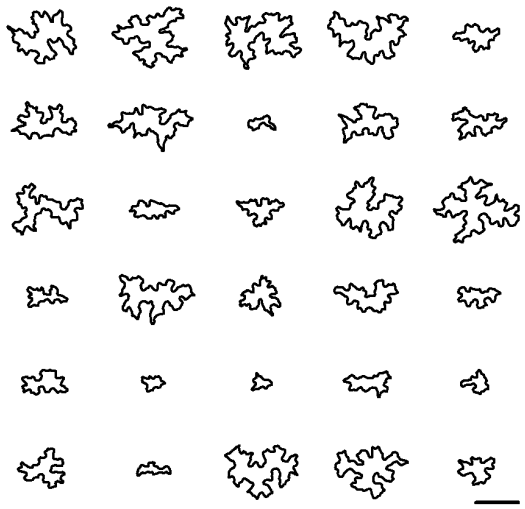


(b) Seedling 2



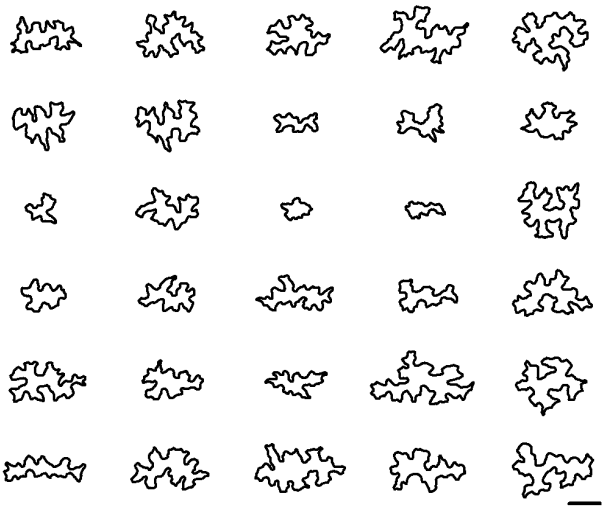
(c) Seedling 3

Figure D.11 – Oy-0



(a) Seedling 1

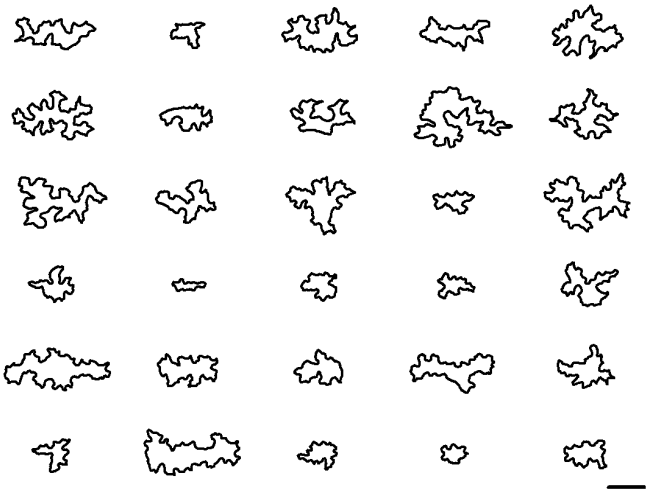
(b) Seedling 2



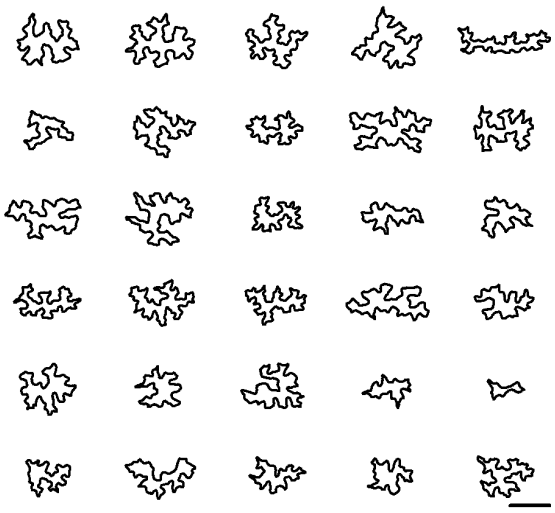
(c) Seedling 3

Figure D.12 – Po-0

(a) Seedling 1

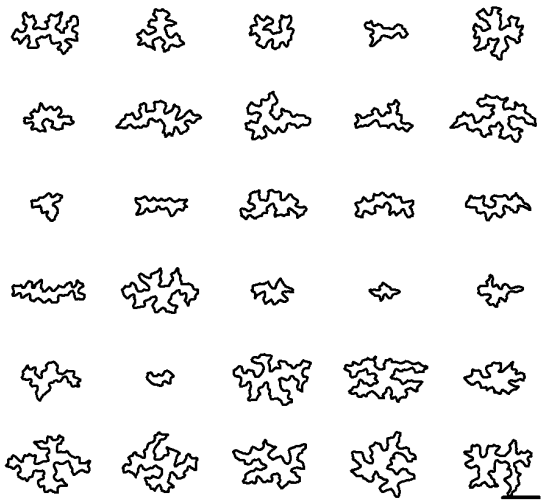


(b) Seedling 2

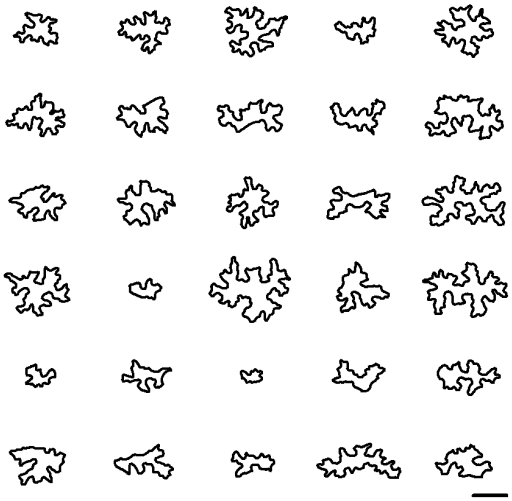


(c) Seedling 3

Figure D.13 – Rsch-4



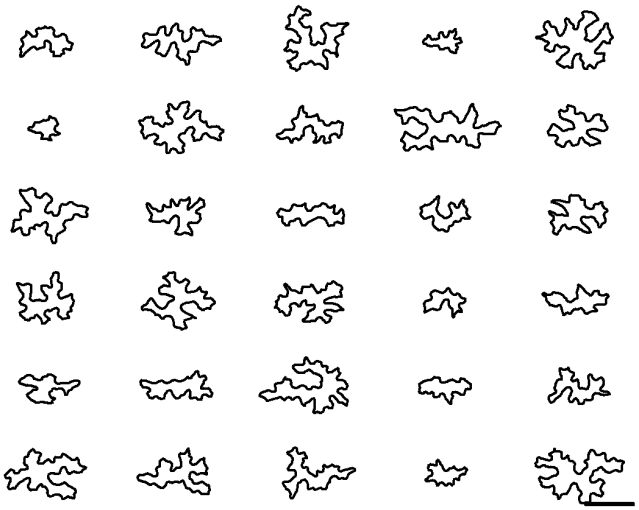
(a) Seedling 1



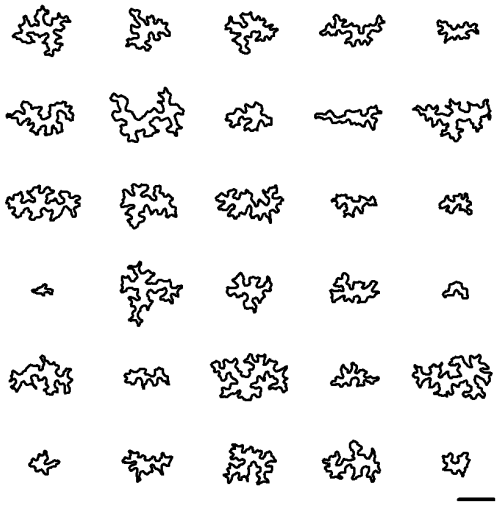
(b) Seedling 2

(c) Seedling 3

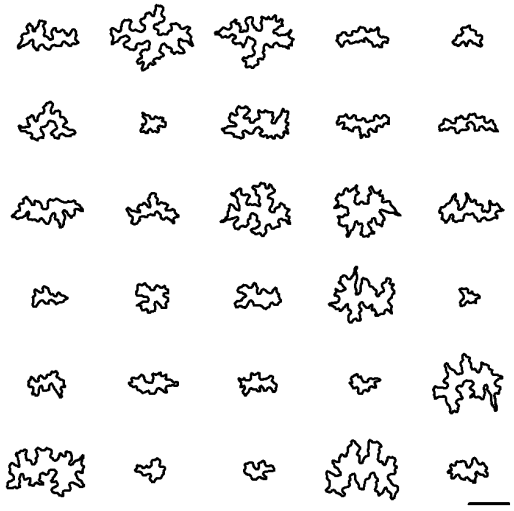
Figure D.14 – Sf-2



(a) Seedling 1

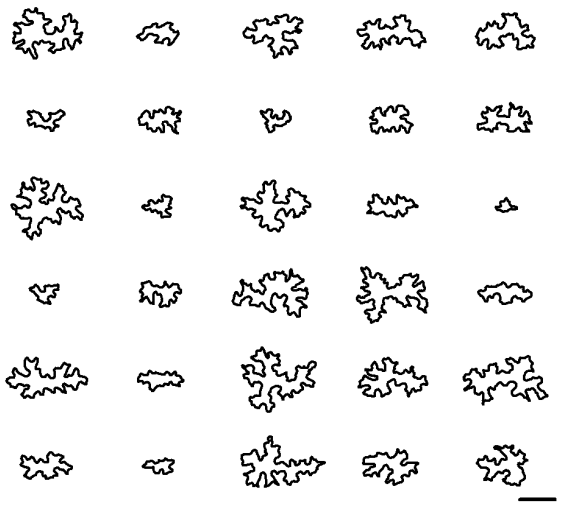


(b) Seedling 2

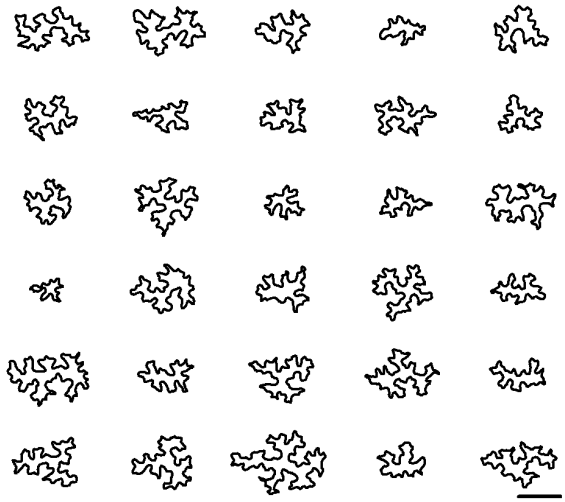


(c) Seedling 3

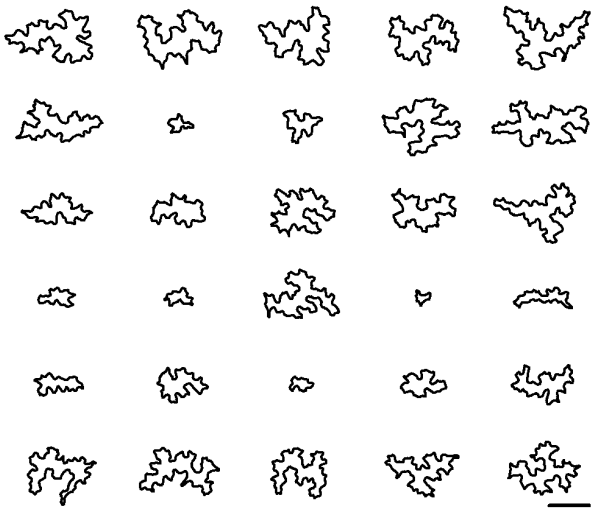
Figure D.15 – Tsu-0



(a) Seedling 1

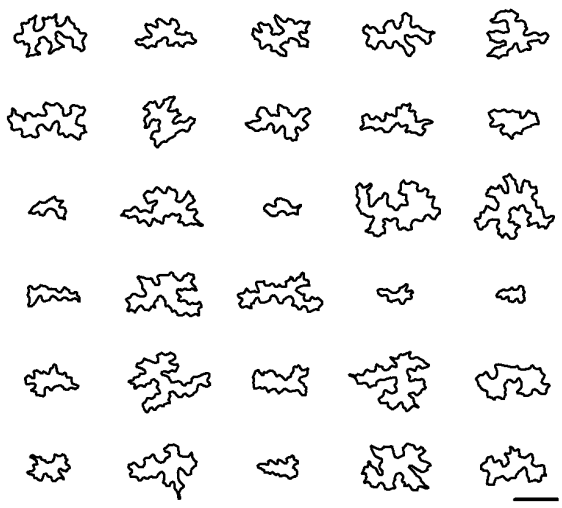


(b) Seedling 2

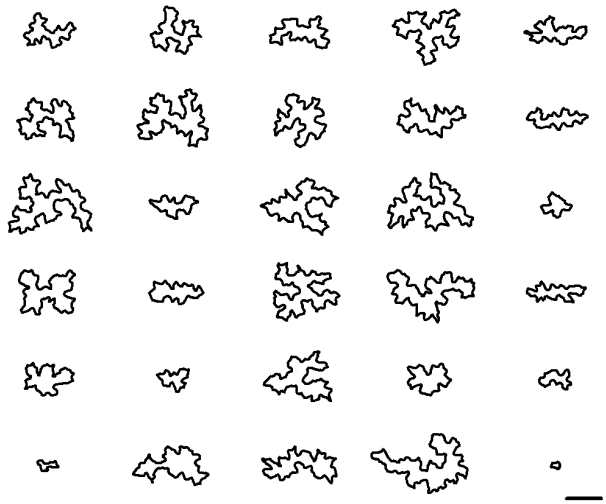


(c) Seedling 3

Figure D.16 – Wil-2



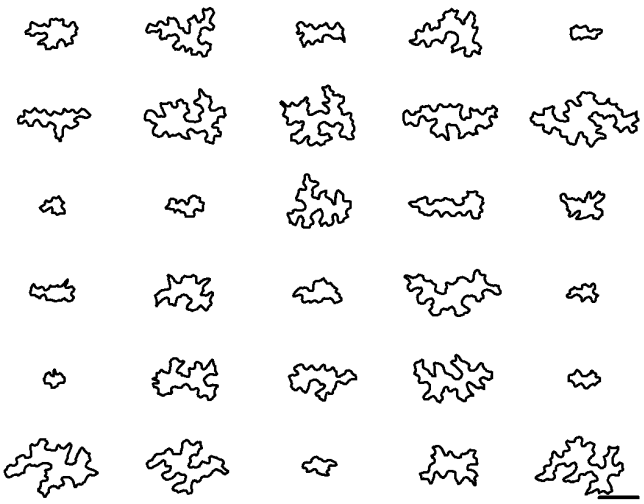
(a) Seedling 1



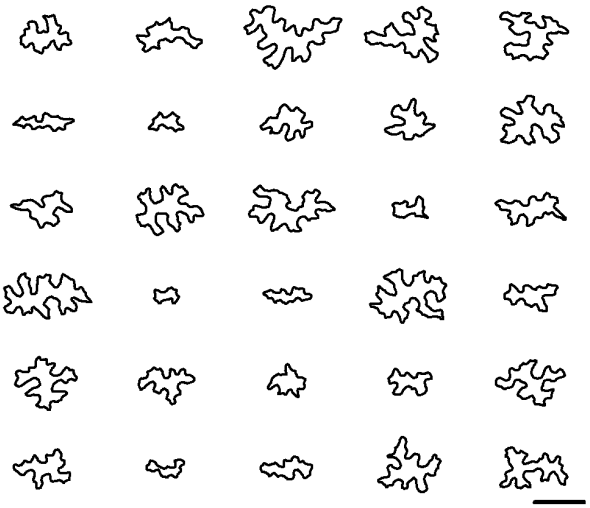
(b) Seedling 2

(c) Seedling 3

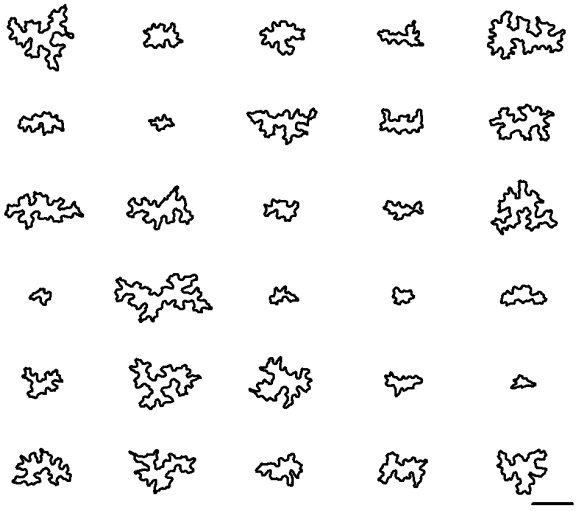
Figure D.17 – Ws-0



(a) Seedling 1

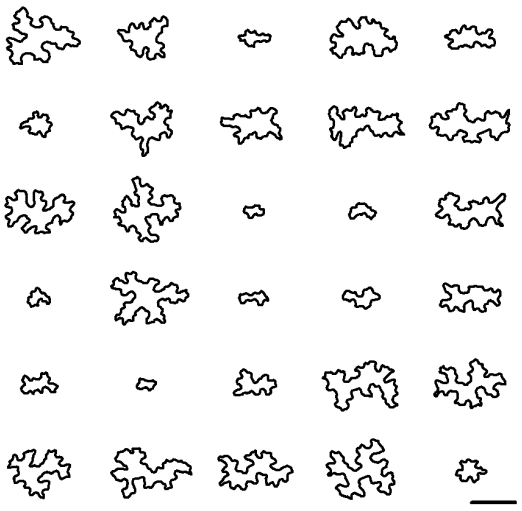


(b) Seedling 2

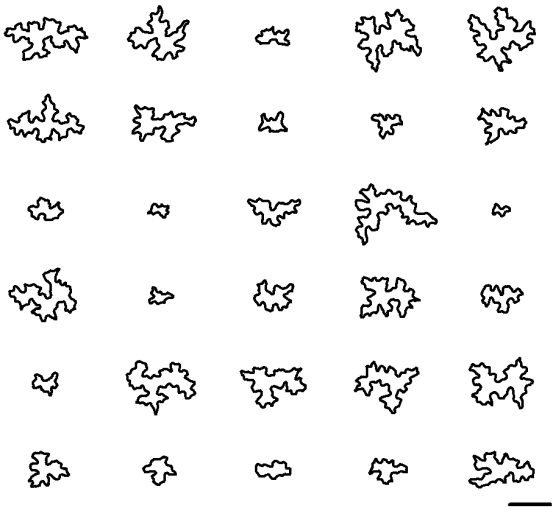


(c) Seedling 3

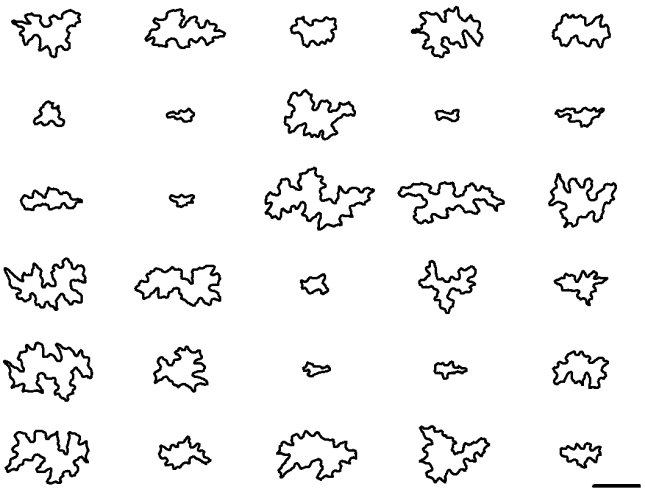
Figure D.18 – Wu-0



(a) Seedling 1

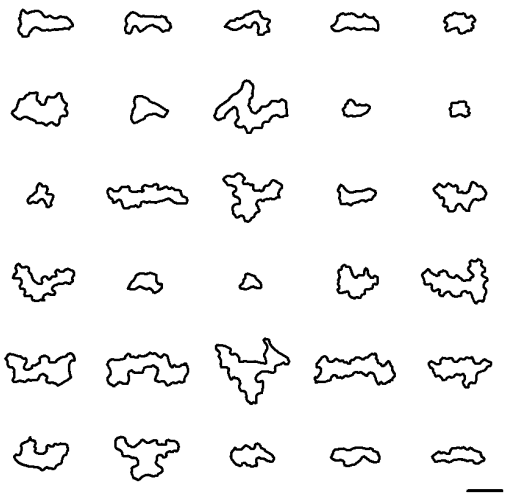


(b) Seedling 2

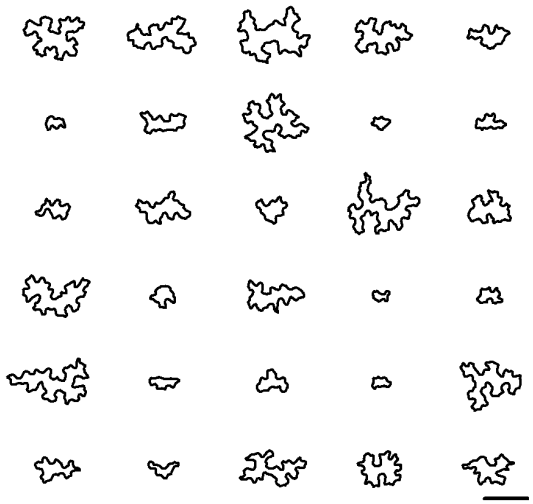


(c) Seedling 3

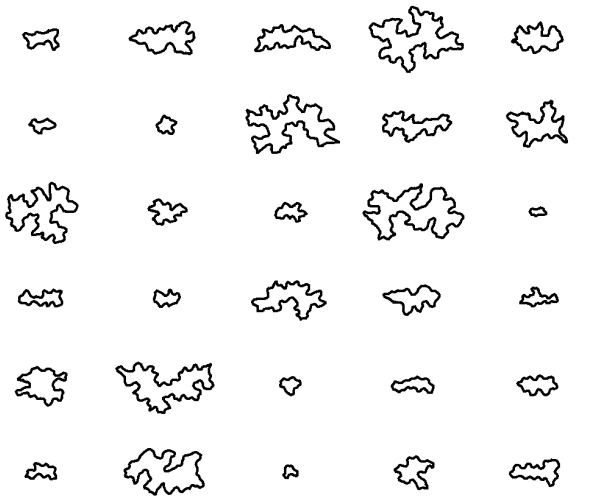
Figure D.19 – Zu-0



(a) *arp3*

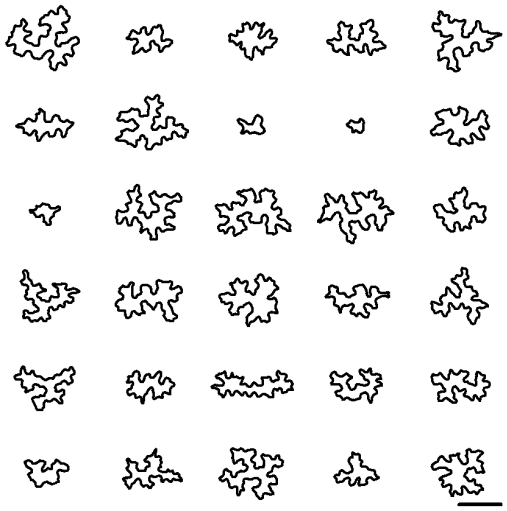


(b) *qua1*

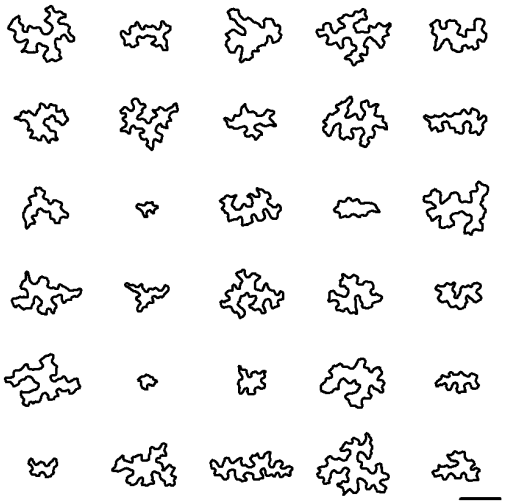


(c) *ric1*

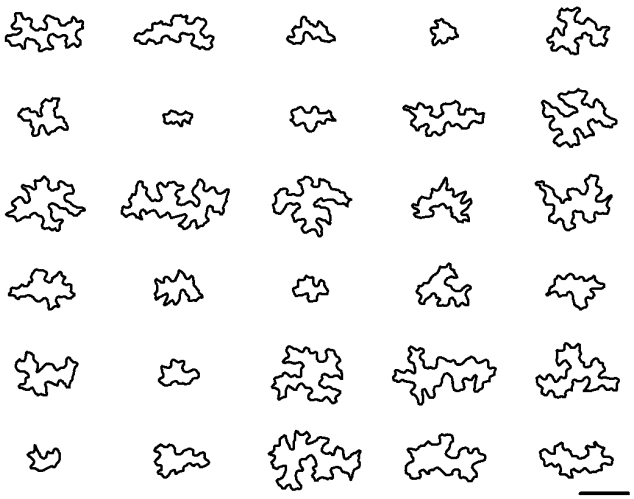
Figure D.20 – Mutant lines I.



(a) *ric4*

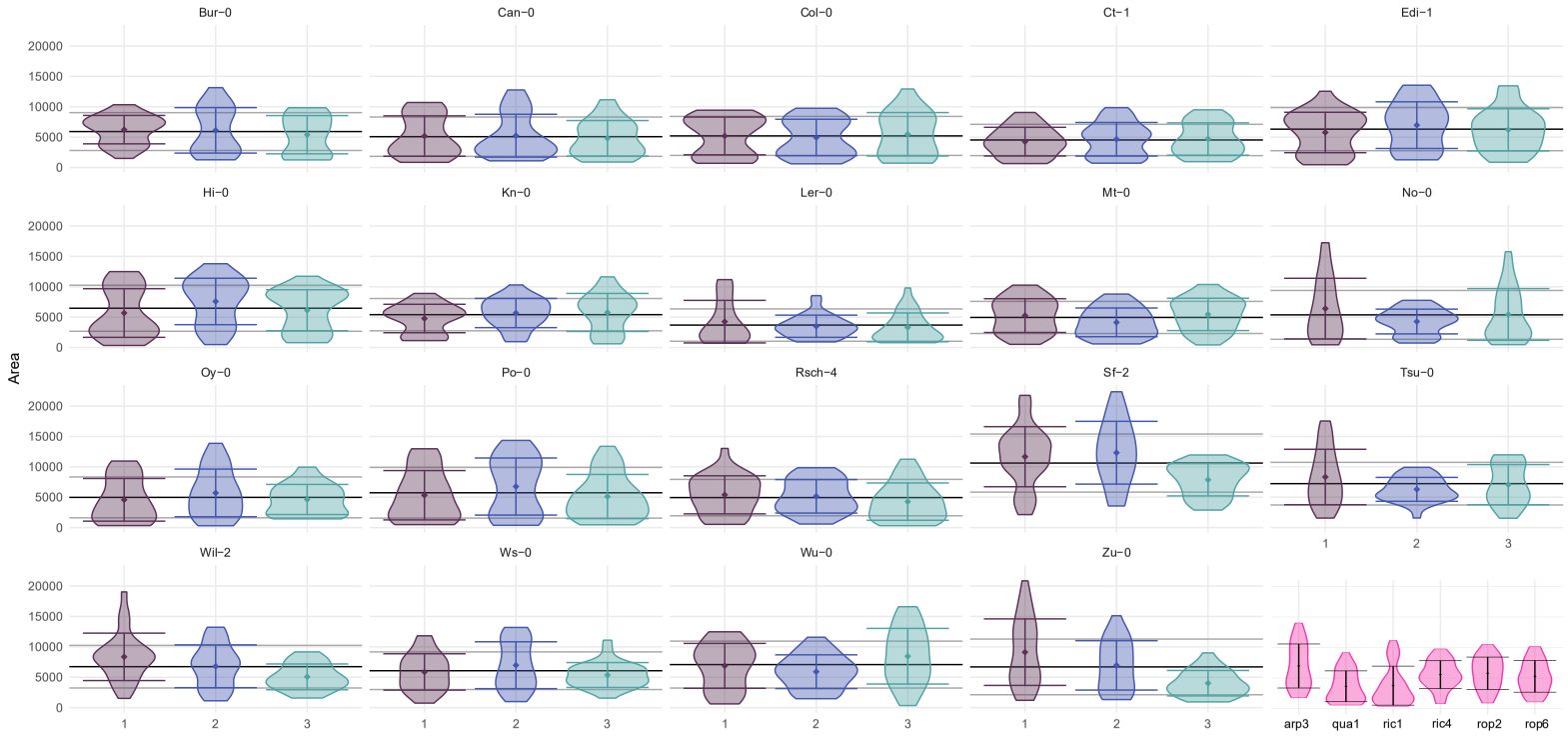


(b) *rop2*

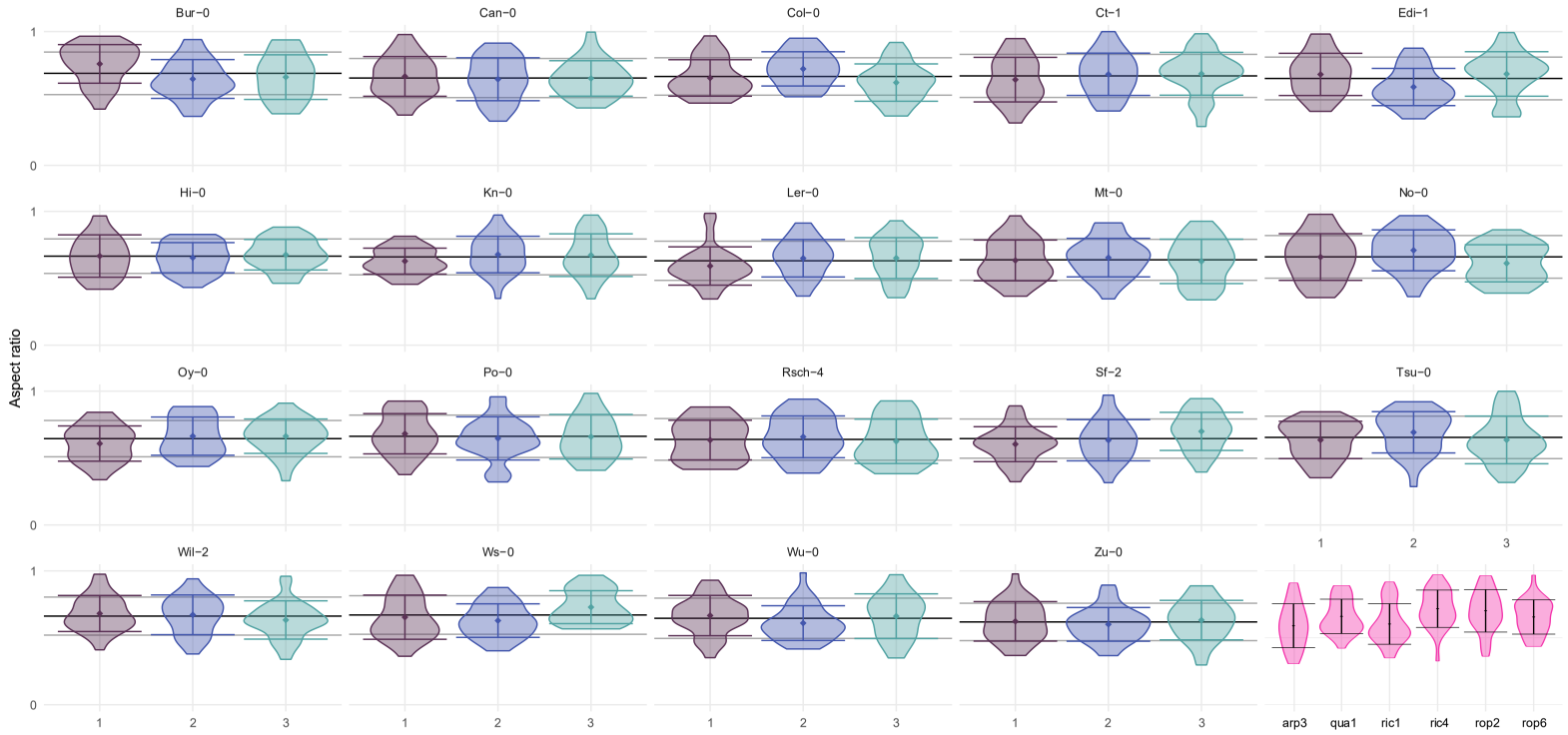


(c) *rop6*

Figure D.21 – Mutant lines II.

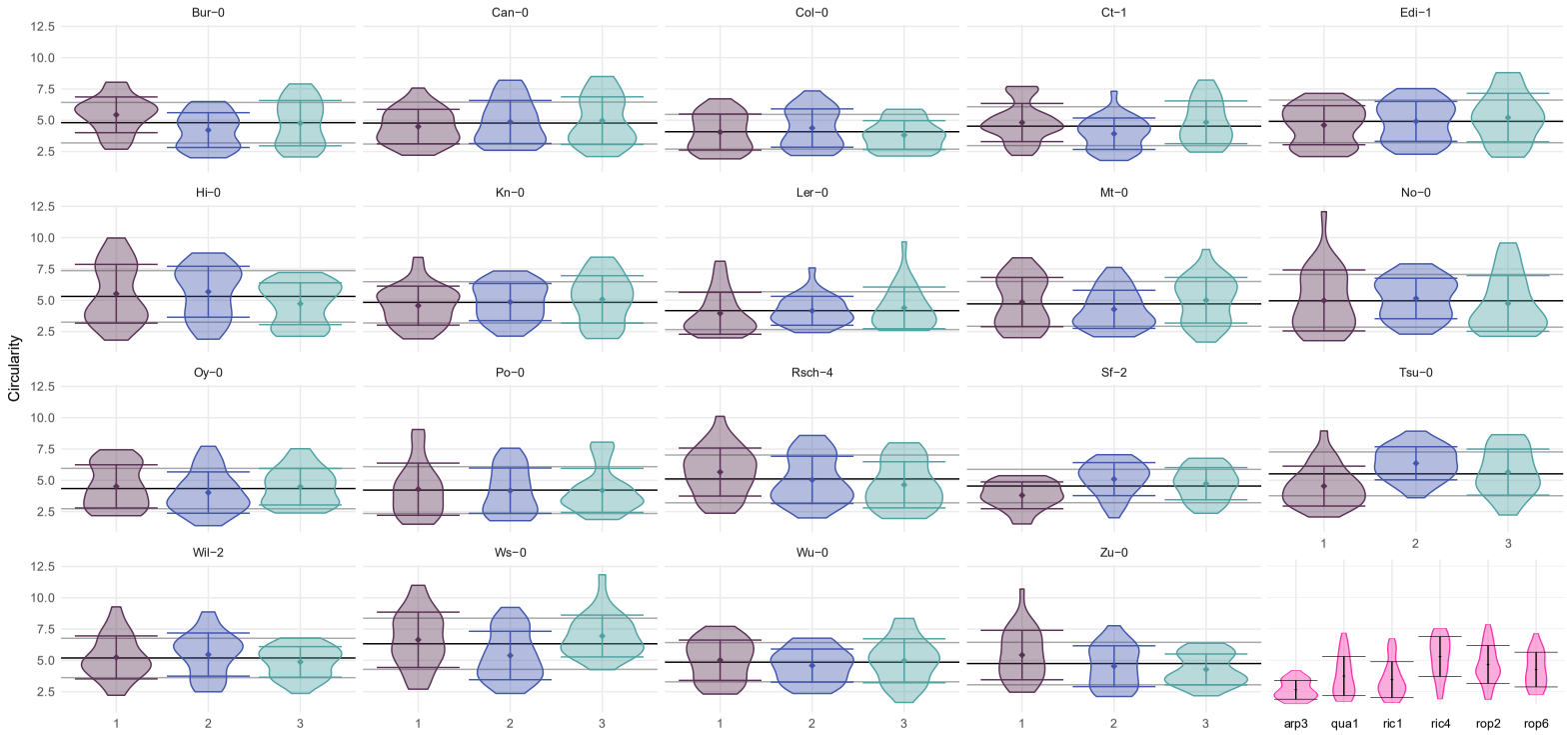


[Area distribution of MAGIC parent and mutant lines]**Area distribution of MAGIC parent and mutant lines** Colours purple, blue and turquoise indicate cotyledons taken from different seedlings for the MAGIC parent lines, pink is used for the various mutant lines. Error bars represent the mean and the standard deviation for the sample. Horizontal lines in the panels for the MAGIC parent lines indicate the mean and mean±standard deviation for the line.



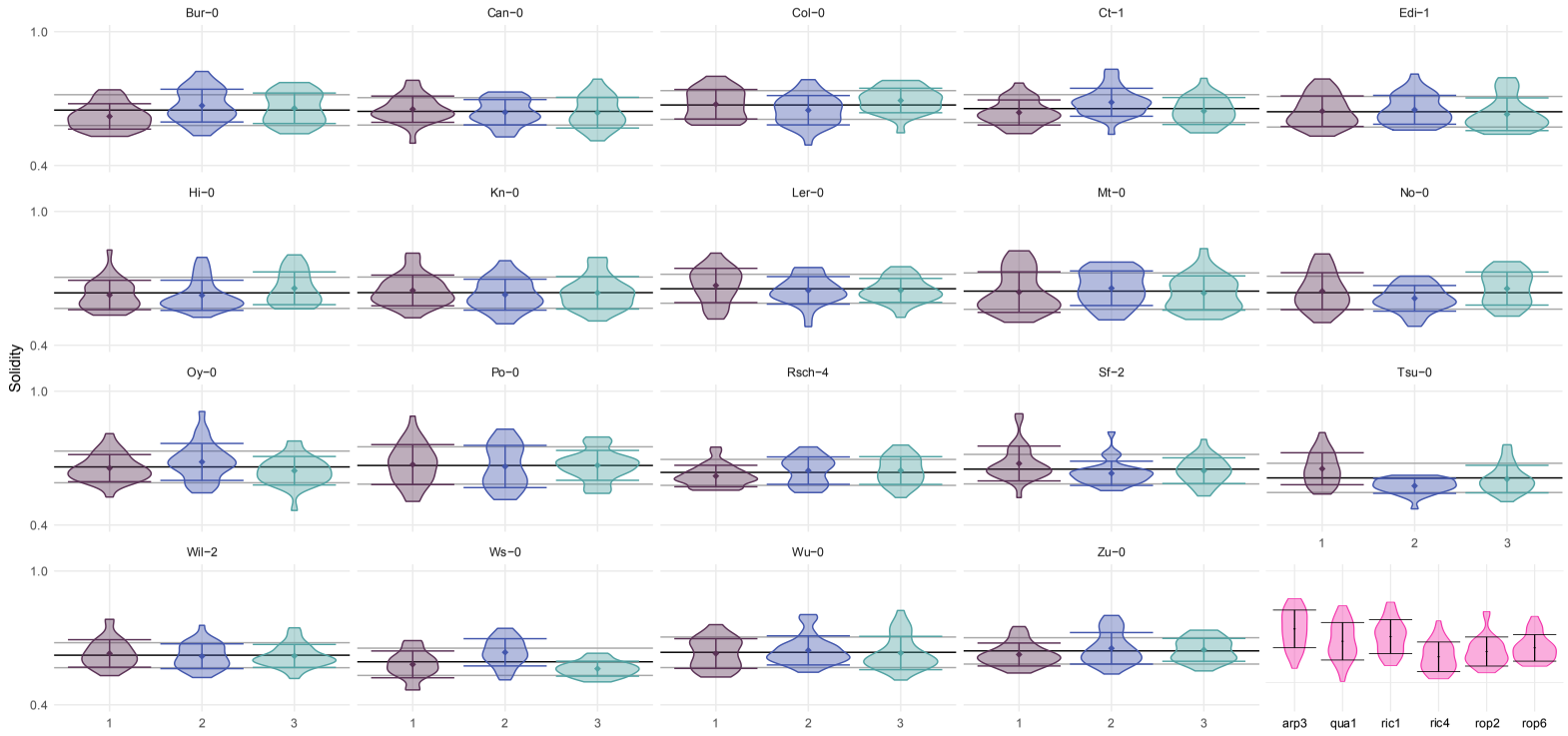
[
Aspect ratio distribution of MAGIC parent and mutant lines]**Aspect ratio distribution of MAGIC parent and mutant lines**

Colours purple, blue and turquoise indicate cotyledons taken from different seedlings for the MAGIC parent lines, pink is used for the various mutant lines. Error bars represent the mean and the standard deviation for the sample. Horizontal lines in the panels for the MAGIC parent lines indicate the mean and mean±standard deviation for the line.



[
Circularity distribution of MAGIC parent and mutant lines]**Circularity distribution of MAGIC parent and mutant lines**

Colours purple, blue and turquoise indicate cotyledons taken from different seedlings for the MAGIC parent lines, pink is used for the various mutant lines. Error bars represent the mean and the standard deviation for the sample. Horizontal lines in the panels for the MAGIC parent lines indicate the mean and mean±standard deviation for the line.



[Solidity distribution of MAGIC parent and mutant lines]**Solidity distribution of MAGIC parent and mutant lines** Colours purple, blue and turquoise indicate cotyledons taken from different seedlings for the MAGIC parent lines, pink is used for the various mutant lines. Error bars represent the mean and the standard deviation for the sample. Horizontal lines in the panels for the MAGIC parent lines indicate the mean and mean±standard deviation for the line.